MONASH University

The NFY transcription factor complex controls nervous system development in *Caenorhabditis elegans*

Pedro Miguel Lopes Moreira

BSc. (Biology), MSc. (Molecular Genetics and Biomedicine)

A thesis submitted for the degree of Doctor of Philosophy at

Monash University in October 2nd 2020

School of Biomedicine, Nursing and Health Science

Monash Biomedicine Discovery Institute

Department of Anatomy and Developmental Biology

Copyright notice

Abstract

Correct establishment of brain architecture during development is a complex process that requires precise control of cell specification, migration, axon outgrowth, and axon guidance. These events are regulated by multiple transcription factors, guidance systems, and environmental cues.

In this study, I explored transcriptional regulation within the *Caenorhabditis elegans* nervous system development at a single-neuron resolution. In an unbiased forward genetic screen, the Nuclear Factor Y transcriptional complex (NFY) was identified as a regulator of PVQ interneuron fate specification. The NFY complex is one of the most common and conserved transcriptional regulators in eukaryotes, responsible for correct body patterning, through Hox genes regulation, and is associated with protein homeostasis in the mammalian brain. The NFY trimeric complex comprises the NFY-A, NFY-B, and NFY-C subunits that regulate gene expression through direct binding to regulatory elements in target gene promoters.

Here, I identified the molecular network through which the NFY complex directly regulates a battery of genes that define the fate of the PVQ interneurons. In addition, I found that the NFY complex controls the expression of pan-neuronal genes in the PVQs and in ventral nerve cord motor neurons. Thereby, the NFY complex plays an important function in regulating both neuron-type specific and pan-neuronal gene signatures.

Declaration

This thesis is an original work of my research and contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

Signature:

Print Name: Pedro Miguel Lopes Moreira

Date: 2nd of October 2020

Acknowledgements

Firstly, I would like to thank my supervisor Professor Roger Pocock for his guidance and support during the last 4 years of this amazing adventure. Roger, thanks for always being around to answer my endless questions, guiding me towards the correct answer, pushing me forward, maintaining a positive view about the project and helping me to develop new skills.

Thanks also to my associate supervisor Dr. Sandeep Gopal and my advisory panel, Professor John Bertram, Assoc. Prof. Robert Bryson-Richardson, and Dr. Luise Cullen-McEwen for their valuable feedback throughout my PhD.

Secondly, to all the members of the Pocock lab, past and present, for providing a great environment to do science. In particular, I would like to thank Steffen Nørgaard for his support and willing to share his expertise during my first year. To Ava Handley and Rasoul Godini for the discussions about science and life, and the feedback about my work. To Oguzhan, Tessa, Wei and Wenyue for being my PhD buddies and making this journey smoother and for being supportive. To Rebecca, Faiz, Leelee, Qi and Chris thanks for helping me and for making life in the lab more interesting.

Finally, a few personal acknowledgements. To my beloved partner Rita, thanks so much for always being there for me and help me with essential feedback about the writing of my thesis. There are no words to thank you enough for all the effort invested. To mum and dad, for always believing in me and for providing the necessary support to take this challenging journey.

Table of contents

List of Tables

Chapter 2. Materials & Methods

8. Appendix

List of Figures

Chapter 1. Introduction

Chapter 3. Characterization of neuron-specific transcriptional elements on the promoter of the *sra-6* gene in C. elegans

Chapter 4. Dissecting the role of the NFY transcription factor complex in nervous system development

5. Endogenously tagging the NFY complex subunits using CRISPR-Cas9 to enable proteomic analysis.

Chapter 6. Final Discussion

List of abbreviations

bHLH – Basic helix-loop-helix

bp – Base pairs

CNS – Central nervous system

CTCF – Corrected total cell fluorescence

C. elegans – Caenorhabditis elegans

DNA – Deoxyribonucleic acid

GFP – Green Fluorescent Protein

GPCR – G protein-coupled receptor

HA – Human influenza hemagglutinin

HFM – Histone-folding motifs

HSP-70 – Heat shock protein 70

HSN – Hermaphrodite specific neurons

IP – Immuno-precipitation

mM – Millimolar

NGM – nematode growth medium

NFY – Nuclear Factor Y

n.s. – Not significant

PCR – Polymerase Chain Reaction

RFP – Red fluorescent protein

SD – Standard deviation

SEM – Standard Error of Mean

TF – Transcription Factor

UTR – Untranslated Region

Ub – ubiquitin

VNC – Ventral nerve cord

YFP – Yellow fluorescent protein

Chapter 1. Introduction

The establishment of correct nervous system architecture during development is a complex process that requires precise genetic control. The neuronal development process involves three major steps: migration of neurons, generation of specific neuronal fates and extension of the projections (axons and dendrites) from the cell body. I will explain these processes in detail with specific reference to the *Caenorhabditis elegans* nervous system. First, neuronal progenitor cells are generated through precisely controlled cell divisions during early embryogenesis (Lee and Goldstein, 2003; Sulston *et al.*, 1983). Next, some neurons need to migrate from the site of generation to their specific functional location in the nervous system. As an example, the hermaphrodite specific neurons (HSN) migrate from the posterior region of the embryo to the ventral mid body region of the embryo during embryogenesis (Desai *et al.*, 1988). Migration of the HSNs is dynamic and requires signals from guidance molecules and the surrounding environment in order to reach the target area. For example, the HSNs are not able to migrate to their correct target area when the worms are under hypoxic stress (Pocock and Hobert, 2008).

Following migration, neuronal progenitor cells subsequently divide and differentiate in order to commit to specific cell lineages that will give rise to most of the cells in the *C. elegans* nervous system. The generation and differentiation of specific neurons is precisely regulated by transcription factors that regulate gene expression within each neuron. For example, a major conserved proneural transcription factor is *atonal*, a basic helix-loop-helix (bHLH) transcription factor encoded by the *lin-32* gene in *C. elegans*. Studies by Bertrand *et al.*, 2002 revealed that expression of this proneural gene drives the generation of neuroblasts at the expense of epidermal cells (named hypodermal cells). This means that the absence of proneural expression causes impairment of neuronal cell fate acquisition. For example, cell lineages that normally give rise to neurons will instead give rise to skin cells (Baumeister *et al.*, 1996; Frank *et al.*, 2003; Zhao and Emmons, 1995).

Once neurons complete their migration to their target destinations and differentiation processes, they are stimulated to extend neurites, which form axons and dendrites that are guided, through precise signalling mechanisms, to connect with other neurons or non-neuronal cells.

Axon guidance, outgrowth, and synapse formation represent the last steps in nervous system development, together with the expression of additional effector genes in the neuron that allow for the functions of the mature neuron. Axon and dendrite outgrowth are essential for the formation of structures like the ventral nerve cord (VNC) in *C. elegans* and spinal cord in humans. In *C. elegans*, the VNC is a structure composed by two asymmetrical bundles of neurons, one on the right composed of more that 50 axons and one on the left side composed of 4 axons (the PVPR, PVQL, AVKR, and hermaphrodite specific neuron left (HSNL)) (White *et al.*, 1976). Correct extension of the VNC axons relies on the expression of guidance signals and adhesion molecules expressed by hypodermal cells and pioneer neurons. Pioneer neurons such the AVG in the right fascicle and the PVPR in the left fascicle are the first neurons to extend and serve as a guide for follower neurons (Durbin, 1987). Both hypodermal and pioneer neurons express a multitude of conserved guidance molecules on their surface or secrete them into the extracellular environment during development. Examples of conserved signalling molecules utilised by the *C. elegans* nervous system to enable neuron navigation include netrin/UNC-6, Ephrin receptor/VAB-1 and the Robo receptor/SAX-3 (Hedgecock *et al.*, 1990; Seeger *et al.*, 1993; Zallen *et al.*, 1999).

Previous studies demonstrated that growth cones, which are structures at the distal tip of outgrowing axons (Figure 1.1), are crucial for sensing and responding to environmental guidance cues (Letourneau *et al.*, 1992; Tessier-Lavigne and Goodman, 1996). These cues generally act in a binomial way, either by attracting the growth cone like UNC-6/Netrin, or by repelling the growth cone from the midline of the VNC like Slit/Robo or Semaphorins (Charron *et al.*, 2003; Zou *et al.*, 2000). It is the combinatorial effect of these signalling molecules that allow axon guidance and outgrowth to occur. In order to sense these molecules, growth cones also express membrane receptors that respond to these molecules. Upon interacting with one of

these guidance molecules, an intercellular signalling cascade is triggered leading to the initiation of a dynamic process of actin-cytoskeleton remodelling to determine outgrowth direction. This process of extension and retraction can be divided in three stages: (1) protrusion, (2) engorgement, and (3) consolidation. In the first (1) stage, actin polymerization regulates the extension of the edges of the growth cone membrane leading to the formation of finger-like structures, called filopodia, upon sensing specific guidance molecules. Subsequently, in the second (2) stage, microtubules transport organelles and vesicles into the peripheral regions of the growth cone leading to the formation of veil-like lamellipodial structures. Finally, in the last (3) stage, the growth cone stabilizes the filopodia and contracts into a cylindrical shaft (Dent *et al.*, 2011; Gallo and Letourneau, 2004; Pak *et al.*, 2008; Zhou and Cohan, 2004). The correct organization and polymerization of actin filaments in filopodia and lamellipodia is a major determinant of growth cone architecture and motility (Figure 1.1) (Gallo and Letourneau 2004), which act in concert with microtubule assembly to drive axon outgrowth.

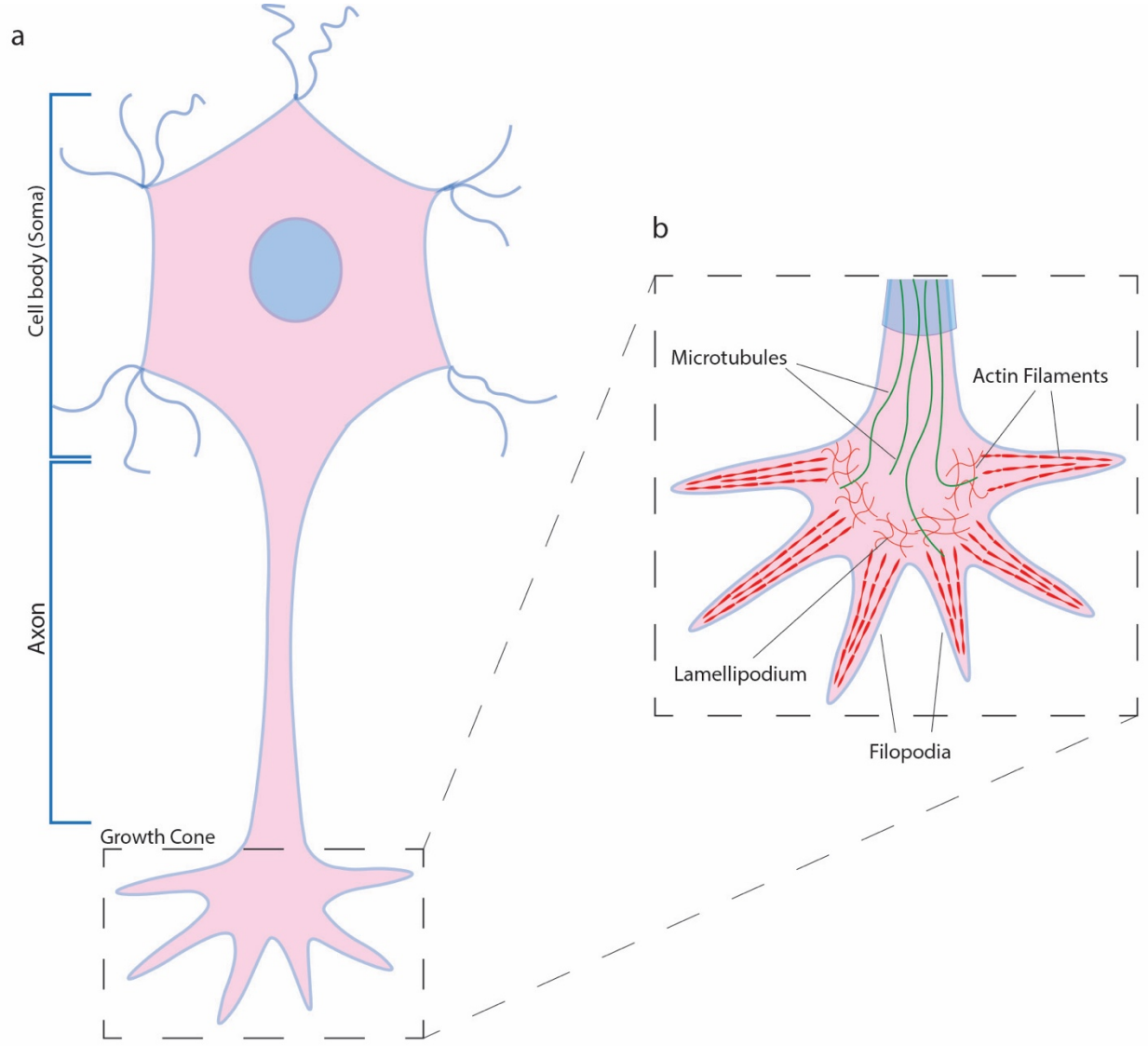

Figure 1.1. Schematic of a neuron and growth cone morphology. A – Overall neuron morphology. **B –** Growth cone morphology. Actin polymerization (in red) is responsible for the protrusion of the filopodia, microtubules (in green) are responsible for transporting vesicles and organelles to the lamellipodium and finally the stabilization of specific filopodia leads to the consolidation of the axon shaft into a cylindrical shape structure.

Lastly, once the nervous system is formed and neurons complete their differentiation and axon extension they need to be maintained in their correct position (Aurelio *et al.*, 2002; Bülow *et al.*, 2004). Previous studies in *C. elegans* revealed that dedicated molecular mechanisms need to occur for the nervous system to maintain its complex architecture after hatching, such as the expression of immunoglobulin-domain proteins and the FGF Receptor (Aurelio *et al.*, 2002; Bülow *et al.*, 2004). As an example,

studies performed by Aurelio *et al.* (2002) showed that immunoglobulin-domain proteins (ZIG-1, ZIG-2, ZIG-3, ZIG-4, ZIG-5 and ZIG-8) were required for the maintenance of the ventral nerve cord organization in *C. elegans*. In this study, Aurelio and colleagues revealed that *zig-4* mutant worms were able to correctly form the PVT interneurons (Aurelio *et al.*, 2002). However, shortly after hatching *zig-4* mutant worms would start to present PVT axon positioning defects (Aurelio *et al.*, 2002). These findings showed that neurons need extracellular cues and adhesion molecules during development and adult life.

Studying the molecules involved in these steps is crucial to build an understanding of pathways involved in nervous system development, as well as the complex mechanism that control its function and dysfunction. In the next sections, I will review how transcriptional regulation controls neuronal development, and describe the molecules that are essential for the correct generation, terminal fate acquisition and extension of neurons.

1.1. Role of transcriptional regulation in neurogenesis

For correct neuronal development to occur, differential gene expression or repression is required in precise spatial and temporal patterns to create molecular programs for neuronal fate acquisition. Previous studies demonstrated that these well-orchestrated programs generate combinatorial molecular cues that provide specific neuronal traits and functions - e.g. expression of specific ion channels or neurotransmitter receptors, on each individual neuronal type and their subsequent survival (Kerk *et al.*, 2017; Rogers *et al.*, 2009; Vidal *et al.*, 2018) (Figure 1.2).

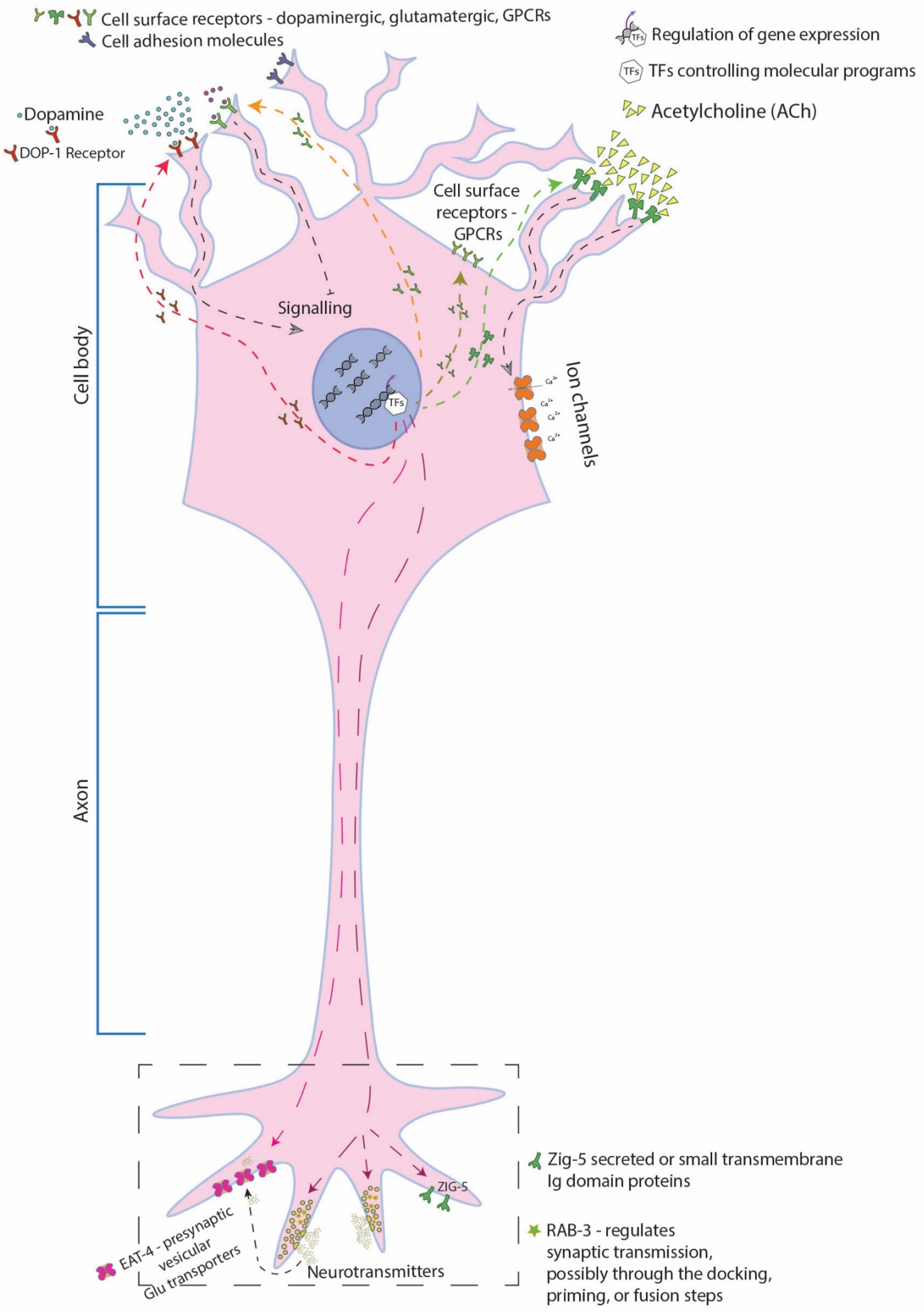

Figure 1.2. Schematic representation of transcriptional regulation during neurogenesis. The image shows dynamic TF networks responsible for acquisition of specific traits that provide neuronal identity – such as GPCRs, neurotransmitters and their receptors.

In recent years, researchers started to grasp the complexity of this molecular mechanism and identified which molecules are required for correct brain development. Currently, it is well known that the key regulators of these programs are proteins called Transcription Factors (TFs) - proteins known to control every one of the steps of neuronal development mentioned above. TFs are commonly defined as proteins with the ability to recognize and bind to short DNA sequences in order to regulate gene expression (Mitchell and Tjian, 1989).

1.1.1. Transcription factor dynamics and function

One of the more challenging aspects in characterising any TF relies on the understanding of how TFs work and what synergies are required for each gene to be expressed. It is well known that some TFs are able to perform different functions, by binding and modulating the expression of different genes, in different cells types or in different stages of development. For instance, studies performed by Jacob *et al.* (2013) revealed that by varying the expression levels of *Ascl1* TF it is possible to induce the generation of different neuronal types in the spinal cord and hindbrain. Moreover, *Ascl1* was also shown to be important at later stages of nervous system formation to control the radial migration of differentiating neurons present in the cortical plate through Rnd-mediated inhibition (Pacary *et al.*, 2011). However, how certain TFs are able to perform different functions in different cells and/or at specific stages of development remains not well-understood.

Currently, it is well accepted that TFs act by controlling a repertoire of genes required for the differentiation of each neuronal type. During nervous system development, ectodermal cells need to express TFs that are members of the *Sox* and *Zic* families in mammals in order for these progenitor's cells to be refractory to BMP and Wnt signals (required to block neuronal induction), avoiding the epidermal fate (Moody *et al.*, 2013; Rogers *et al.*, 2009; Sasai, 1998). It is hypothesized that these TFs function as modulators of neuronal cell competence, since they confer the ability for cells to ignore specific signals (Moody *et al.*, 2013; Rogers *et al.*, 2009; Sasai, 1998). *Sox* and

Zic TFs families present overlapping expression domains and are required to work together to maintain precise balance between proliferation and differentiation of neural ectodermal precursors (Moody *et al.*, 2013; Rogers *et al.*, 2009; Sasai, 1998). As an example, SoxB1 and SoxB2 subfamilies have been suggested to have opposing roles (Sandberg *et al.*, 2005). Previous studies revealed that SoxB1 subfamily repression leads to a decrease in d1-crystallin enhancer DC5 and promotes differentiation of precursor cells (Uchikawa *et al.*, 1999), while repression of SoxB2 subfamily leads to precursor cell proliferation and impaired neuronal differentiation (Uchikawa *et al.*, 1999). Based on these findings, it is thought that SoxB1 and SoxB2 subfamilies balance each other for correct nervous system development to occur. Another example is the mutant mice for bHLH (*Ebf1* gene), which shows incorrect migration of motor neurons to rhombomere 5 and rhombomere 6 during hindbrain development. The migration defect occurs due to transcriptional dysregulation of the cell membrane protein *TAG-1,* and the cell-adhesion molecule *cadherin 8* in the facial branchiomotor neurons (Garel *et al.*, 2000). Studies performed in *Ebf1* mutant mice revealed that facial branchiomotor neurons exhibit a reduction of expression of *TAG-1* in rhombomere 5 and an increase in the expression of *cadherin-8 on* rhombomere 4 and 5 (Garel *et al.*, 2000). Changes in the spatial and temporal molecular signature of these neurons impair their ability to respond to environmental cues presented in rhombomere 5 and 6, leading to incorrect migration (Garel *et al.*, 2000). Furthermore, previous studies in *C. elegans* showed that neuronal fate programs rely on TFs, defined as terminal selectors, that unlike Ebf1 and Sox act at later stages to achieve the unique molecular signature of each neuron (Hobert, 2011). Therefore, TFs are responsible for the molecular network regulation required for sensory receptor, neurotransmitter receptor, ion channel, adhesion molecule and neuropeptide formation (Hobert, 2011). For example, the ETS factor *ast-1* is essential for the control of all dopamine pathway genes (Flames and Hobert, 2009). Accordingly, mutation analysis of *ast-1* revealed a failure in dopaminergic terminal fate acquisition (Flames and Hobert, 2009). Additionally, genetic studies performed by Masoudi *et al.* (2018) revealed that AS-C homolog *hlh-4*, a bHLH TF, is responsible for the molecular signature transcriptional

regulation of ADL neurons in *C. elegans* by controlling neuronal trait formation, such as GPCRs, glutamine receptors, neuropeptides, innexins and Ca^{2+} exchangers.

Overall, the same TF is able to interact with different proteins, defined as co-activators or repressors of transcription, to regulate a wide range of processes, such as cell generation, proliferation, differentiation, and maintenance. This characteristic is the key element that underlines how these regulatory TF networks are dynamic within different cells of an organism (Gertz *et al.*, 2012).

1.1.2. Transcription factor binding motifs

A challenging aspect of TF characterization relies on the identification of DNA binding motifs. These short DNA sequences (5 to 25bp) are conserved and highly degenerated motifs, such as TATA, GC and CCAAT boxes, usually located in the promoter region (upstream of any gene), and other domains that are located in more distal regions upstream of the ATG that are able to enhance or repress gene expression through the biding of enhancers/repressors proteins (Peng *et al.*, 2006; Tjian and Maniatis, 1994). Therefore, to correctly characterize TF function, extensive knowledge is required about motif recognition.

Given that the interaction between TFs and DNA is crucial, many binding motifs and corresponding binding sites in TFs are conserved across eukaryotes. To date, there has been strong investment in database construction that include all possible binding motifs using a variety of *in vitro* or *in vivo* methods (Herrero *et al.*, 2016; Jolma and Taipale, 2011). Some of these methods, such as protein binding microarrays and yeast one-hybrid (Johnson *et al.*, 2007; Reece-Hoyes and Marian Walhout, 2012), are able to produce high-throughput data, while others like DNA affinity purification sequencing (DAP-seq) and microfluidics-based mechanically induced trapping of molecular interactions (MITOMI) are only able to produce medium-throughput data (Brooks *et al.*, 2019; Maerkl and Quake, 2007). However, these techniques can accurately discover novel binding sites present on the DNA. In contrast, *in vitro* methods such

Electrophoretic Mobility Shift Assay (EMSA) (Fried and Crothers, 1981), DNA footprinting (Schmitz and Galas, 1979), Isothermal Titration Calorimetry (ITC) (Poon, 2010) and surface plasmon resonance (SPR) (Ritzefeld and Sewald, 2012) are more straightforward techniques that only allow validation of known TF binding sites. In more recent years, the introduction of CRISPR/Cas9 genetic engineering technology led to a shift to an *in vivo* approach. It is now possible to endogenously tag TFs and use chromatin immunoprecipitation sequencing (ChIP) based techniques to identify specific binding sites (Johnson *et al.*, 2007). All the information generated by these methods is compiled and stored in different databases, such as JASPAR (Mathelier *et al.*, 2016), TRANSFAC (Matys *et al.*, 2006) and UNIPROBE (Hume *et al.*, 2015), for further *in silico* studies. However, the inability to precise distinguish between direct and indirect binding is still a limitation. Studies performed in *C. elegans* aimed to fill this gap by using advanced genetics to delete predicted binding sites of promoter regions in order to decipher if a specific TF directly regulates specific downstream genes. This technique uses available TF binding site information, previously gathered by other techniques, to mutagenize specific motifs in the promoter region of the reporter plasmid. Next, this plasmid would be injected into wild-type worms to analyse any change in fluorescence expression. As an example, studies performed by Kratsios *et al.* (2017) aimed to prove that LIN-39, an Hox protein, was responsible for directly controlling the expression of *unc-129* and *del-1* genes known to be involved in motor neuron development. By mutagenizing the LIN-39 putative binding sites in GFP reporters for both genes Kratsios *et al.* (2017) found a reduction in GFP expression that phenocopied the *lin-39* mutant worms. Together, these results suggested that LIN-39 directly regulates *unc-129* and *del-1* in motor neurons (Kratsios *et al.*, 2017).

While determining whether the binding site is present is the first step to understand the molecular mechanism associated with a TF, further analyses such protein-protein interaction, TF domain deletion, are often necessary to understand the TF mechanism of action.

1.2. NFY TRANSCRIPTION FACTOR

One of the TFs known to have a role in neuronal maintenance is the Nuclear Factor Y (NFY) transcription factor family that shows a high affinity for the CCAAT motif in eukaryotes (Tjian and Maniatis, 1994). The NFY complex is one of the most common and conserved transcription factors in eukaryotes. It is known to regulate the expression of genes associated with nervous system development and maintenance, and body patterning regulation (Deng *et al.*, 2007; Yamanaka *et al.*, 2014; Yamanaka *et al.*, 2016). In *C. elegans,* the NFY family regulates the expression of the Hox gene *egl-5* and control tail patterning. Additionally, in Xenopus and mammals, the NFY family is involved in heat shock protein 70 (HSP-70) expression regulation (Arents and Moudrianakis, 1995; Benatti *et al.*, 2016; Deng *et al.*, 2007; Li *et al.*, 1992; Mantovani, 1999; Yamanaka *et al.*, 2008). More recently, Yamanaka and colleagues demonstrated that the NFY complex is also important for neuronal maintenance in the mammalian brain (Yamanaka *et al.*, 2016).

1.2.1 Overall structure of the NFY complex and binding to DNA

Studies in yeast showed that the NFY complex is a heterotrimeric protein composed of three different subunits, NFYA-I, NFYB-I and NFYC-I, which interact with each other in a specific manner through evolutionary conserved motifs called HAP2, HAP3 and HAP5 (Baxevanis *et al.*, 1995; Mantovani, 1999). HAP3 and HAP5 possess conserved histone-folding motifs (HFM) that are composed by three or four α-helices (α1, α2, α3 and αN/C) separated by two loops (L1 and L2) (Baxevanis *et al.*, 1995). Functional studies showed that these conserved HFMs interact with each other in a head to tail sequence, with α1 from NFYC interacting with α3 from the NFYB and the α3 from NFYC interacting with α1 from the NFYB, to form dimers (Figure 1.3). The NFYC/NFYB dimer forms a complex surface that enables NFYA-1 recruitment to form a heterotrimeric complex with a high affinity for CCAAT boxes (Figure 1.3) (Mantovani, 1999).

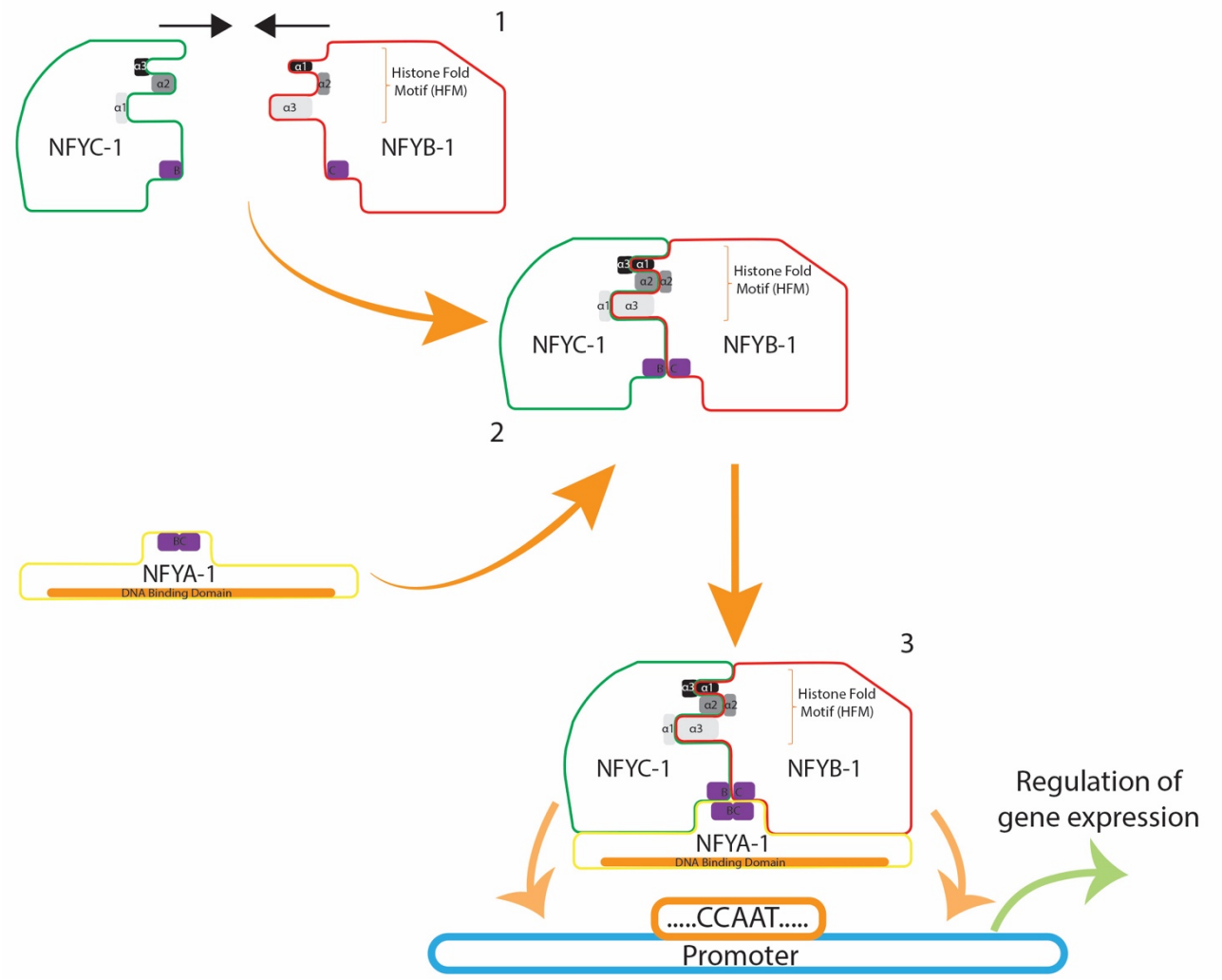

Figure 1.3. Schematic of NFY transcription factor complex formation. The NFY transcription factor is composed by three different subunits called NFYA-1, NFYB-1 and NFYC-1. Together they form a heterotrimeric protein that binds the promoter of genes to regulate gene expression in conjunction with coactivators. **(1)** In first step of complex formation, the NFYB and NFYC subunits bind through conserved HFM regions to form a dimeric protein. **(2)** The heterodimeric (NFYC/NFYB) protein recruits the NFYA subunit to enable DNA binding. **(3)** Finally, the DNA binding domain of the NFYA subunit recognizes CCAAT boxes in the promoter of target genes to control gene expression.

Notably, the NFYA subunit possesses a DNA binding domain that is able to recognize specific regions in target gene promoters– a consensus CCAAT box usually located in the minor groove of DNA, which is one of the most frequent and conserved regulatory motifs present in the promoter regions in eukaryotes (Figure 1.3) (Deng *et al.*, 2007; Mantovani, 1998). Studies performed by Nardini *et al.* (2013) revealed that the NFYA-1 A2 helix is a key element for correct binding to DNA. This key element recognizes

and binds to the CCAAT boxes present in the minor groove of DNA creating a shift in the DNA structure. This architectural change promotes a 48-degree positive roll centred on the CCAAT box that will widen the DNA minor groove while maintaining major groove stability (Nardini *et al.*, 2013). Further genetic studies revealed that specific residues, such as Arg274 (NFYA) – O2 atom of C (13), Arg281 (NFYA-1) to N3 atom of A 910), are essential for the NFYA-1 function and that any mutation in these residues abolishes DNA binding and results in impairment of the NFY complex function (Nardini *et al.*, 2013; Oldfield *et al.*, 2014).

Genetic studies revealed that the NFY complex requires all five nucleotides of the CCAAT box for correct binding (Deng *et al.*, 2007). Mutations in any of the CCAAT nucleotides abolishes binding and function of the NFY complex (Deng *et al.*, 2007). CCAAT binding motifs are frequently found -60 to -100 bp from the transcriptional start site and they can be found in 64% of the human genes (Mantovani, 1998; Suzuki *et al.*, 2001). Studies by Mantovani (1998) revealed that the CCAAT orientation for these conserved motifs is most prevalent in the promoter regions, when compared with the ATTGG orientation, and that these conserved motifs tend to be located in close proximity to the TATA box. Therefore, it has been hypothesized that in the absence of TATA Binding Protein-TATA box interaction, the NFY complex may function as a conduit between the upstream activators and the general transcriptional machinery (Mantovani, 1998, 1999). Moreover, previous studies performed in mouse embryonic stem cells by Oldfield *et al.* (2014) revealed that NFY complex binding to CCAAT sites is essential to provide chromatin accessibility for other TFs to perform their function, such as Oct4.

Despite progress made in understanding NFY function, the NFY complex's precise role in gene regulation remains largely unknown. So far, NFY complex function has been described as a binding facilitator for other TFs, a co-activator of transcription, and a possible chromatin remodelling agent.

1.2.2. Role of NFYB and NFYC in DNA binding and their function

The role of the NFYB and NFYC subunits in the NFY complex has not yet been fully investigated. Crystallography studies performed by Romier *et al.* (2003) and functional studies performed by Nardini *et al.* (2013) compared the NFYB/NYFC dimer and the H2B/H2A dimer core histones family, and suggested that both NFYB/C and H2B/H2A dimer complex have highly similar structures and binding modes with the α1-α1 regions, while the L1-L2 regions are required for NFY/DNA interaction establishment. These similarities in DNA binding were suggested by studies showing that replacement of the α1 and mutations in specific residues of NFYB-1 (Asn61) and NFYC-1 (Arg47 and Lys49) subunits are sufficient to abolish the NFY/DNA interaction (Zemzoumi *et al.*, 1999). Overall, these studies demonstrated that HFM integrity is essential for dimer formation, NFYA association, and CCAAT-binding. However, the individual function of each subunit is not clear.

Currently, it is thought that both NFYA and NFYC may play a more prominent role in the functionality of the NFY complex when compared with the NFYB subunit. This hypothesis is based in the fact that both of these subunits possess two large glutamine rich domains, over 150 amino acids in length, responsible for the NFY activation (Mantovani, 1998). In support of this theory, removal of the NFYA-1 Q-rich domain leads to the generation of a dominant negative mutant protein that affects the NFY transcription factor activity (Hu and Maity, 2000).

1.2.3. NFY complex interactors (co-activator or repressor of transcription)

One of the key steps to understand how a transcription factor acts is to investigate the possible interaction that may occur with other proteins in transcriptional regulation. It is known that the NFY complex is able to interact with other co-activators/repressors to repress or enhance transcription. These co-activators/repressors are usually able to recognize the motifs flanking the CCAAT boxes in different promoters (Figure 1.4) (Li *et al.*, 1998; Luo *et al.*, 2008).

To date, a few examples of the interaction between the NFY complex and other proteins have been reported. Of importance, previous studies reported that NFY cooperates with the MES-2/MES-6 complex, a Polycomb group of proteins (EDD family), to regulate *egl-5* expression in *C. elegans* (Deng *et al.*, 2007; Mantovani, 1999). In particular, Deng *et al.* (2007) demonstrated that the MES-2/MES-6 complex is able to bind to the NFYA-1 subunits to repress the expression of the *egl-5* gene that is required for the correct tail patterning in *C. elegans*. Another example is the interaction between p300 and the NFY complex in *Xenopus*. Studies performed by Li *et al.* (1998) showed that p300, a coactivator/histone acetyltransferase known to promote chromatin disruption, interacts with the NFYB-1 subunit to regulate *hsp-70* expression (Figure 1.4). Additionally, deletion mutants revealed that the p300 C-terminal domain is responsible for the interaction and acetylation of the NFYB-1 subunit *in vitro* and *in vivo,* contributing for the correct transcriptional regulation of the *hsp-70* gene.

Additionally, the NFY complex has been shown to interact with ATF6α and work as activator of gene expression (Luo *et al.*, 2008). ATF6α belongs to the bZIP TF family, and is known to bind to the CCACG boxes located on the promoter element – Endoplasmic Reticulum Stress Element (ERSE). The ERSE is typically composed of one CCAAT and one CCACG box that are required for the NFY and ATF6 binding (Figure 1.4) (Luo *et al.*, 2008). In agreement with this, studies by Luo *et al.* (2008) revealed that any mutation in the CCAAT box is sufficient to disrupt the interaction between NFY and ATF6, and thus impair gene expression during ER stress.

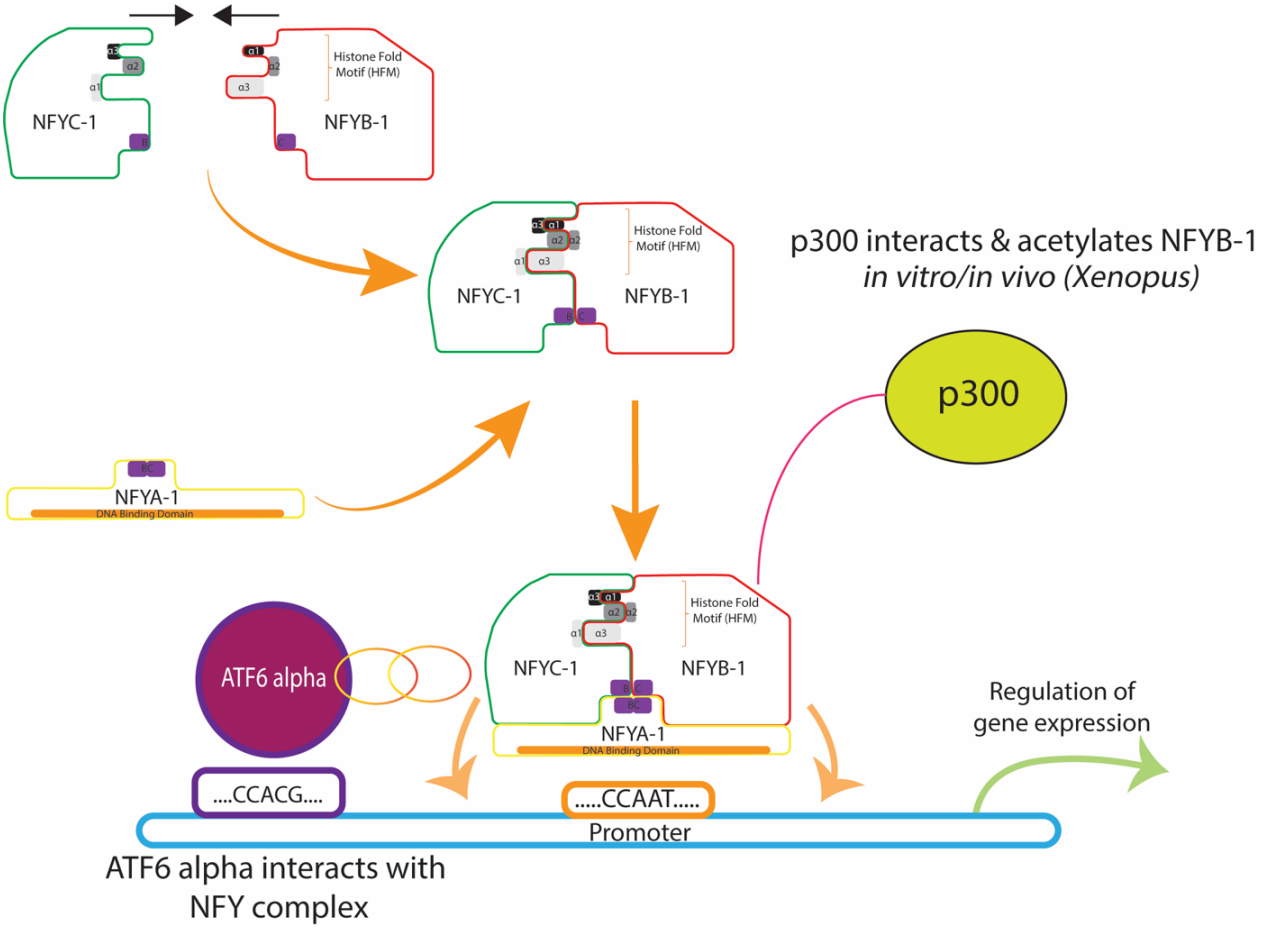

Figure 1.4. Representation of NFY complex and co-activators/repressors interaction. This image shows two examples of known interaction between the NFY complex and other coactivators/corepressors of transcription. On the right, we have depicted the interaction between the NFYB-1 subunit and the p300 histone acetyltransferase. On the left, we have the interaction between ATF6, bZIP TF family, and the NFY complex during ER stress. The double ring structure signifies that it is not known which NFY subunit interacts with a particular protein.

1.2.4. Role of the NFY complex in brain development and disease

The role of the NFY complex during brain development remains largely uncharacterised. One possible explanation for this lack of information is the essential requirement for the NFY complex in most organisms, and as such mutations in any NFY complex subunits causes embryonic lethality (Yamanaka *et al.*, 2014; Yoshioka *et al.*, 2008). To date, most studies performed in eukaryotes require the use of tissue specific knockouts or the use of knockdown techniques, such RNA interference. *C.*

elegans is an excellent model for the study of NFY complex biology as loss of *nfy* genes does not cause lethality (Yamanaka *et al.*, 2014).

In recent years, it was found that the NFY complex is involved in R7 axon pathfinding and correct synapse targeting in *Drosophila* eye development. Studies performed by Morey *et al.* (2008) revealed that mutation in the *nfyc*-1 gene lead to the de-repression of R8-specific transcription factor Senseless3 (*sens*) during R7 neuronal differentiation. These events lead to the R7 axons incorrect targeting of the M3 region, region targeted by R8 axons, leading to incorrect synapse formation (Morey *et al.*, 2008). Moreover, evidence has emerged linking the NFY complex to the nervous system due to the use of tissue specific knockouts in mice. In 2008, it was found that the interaction between the NFY complex and the *hsp-70* gene promoter, which is known to help reduce the toxicity of polyglutamine stretches in Huntington's disease, can be disrupted due to an interaction between mutant Huntingtin protein (Htt) and the polyQ region present in the NFYA and NFYC subunits (Yamanaka *et al.*, 2008). The study revealed that the NFYA and NFYC subunits are sequestered by Htt, both *in vitro* and *in vivo* in a Htt mutant mouse model, promoting transcriptional dysregulation and triggering neurodegenerative disease (Yamanaka *et al.*, 2008). More recently, two studies performed by Yamanaka and colleagues suggested a possible link between the NFY complex and neurodegeneration in a mouse model (Yamanaka *et al.*, 2014; Yamanaka *et al.*, 2016). Specifically, Yamanaka *et al.* (2014) revealed that NFYA-1 knockdown in the forebrain results in cytoplasmic accumulation of ubiquitin (Ub) and Scaffold protein p62, and downregulation of GRP94 and ER chaperones in pyramidal neurons. Subsequently, the authors investigated the connection between the NFY complex and neurodegeneration in motor neurons. Yamanaka *et al.* (2016) showed that NFY complex inactivation, by downregulation of NFYA-1, results in motor neuron degeneration with downregulation of Grp94 and Grp78 but absence of p62 and Ub accumulation. These results suggested that NFY complex inactivation is responsible for the neuronal pathologies (Yamanaka *et al.*, 2016).

However, despite these findings, the function of the NFY complex in the generation and specification of different neuronal types remains unknown.

1.3. *Caenorhabditis elegans* as a model

To better understand how neurons are generated and develop, I am studying neuronal development at single-cell resolution using the *in vivo C. elegans* model. The main advantages of the *C. elegans* are its small size (1 mm in length), transparent body, and completely mapped small nervous system – comprising 302 neurons in the hermaphrodite and 383 neurons in the male (White *et al.*, 1986) (Figure 1.5). Additionally, the fact that the *C. elegans* genome is completely sequenced and *C. elegans* possesses a short life cycle (develops from egg to a young adult worm capable of laying eggs in 47 hours) makes this model an excellent tool to use in forward and reverse genetic screens to identify new genes involved in brain development (Figure 1.6).

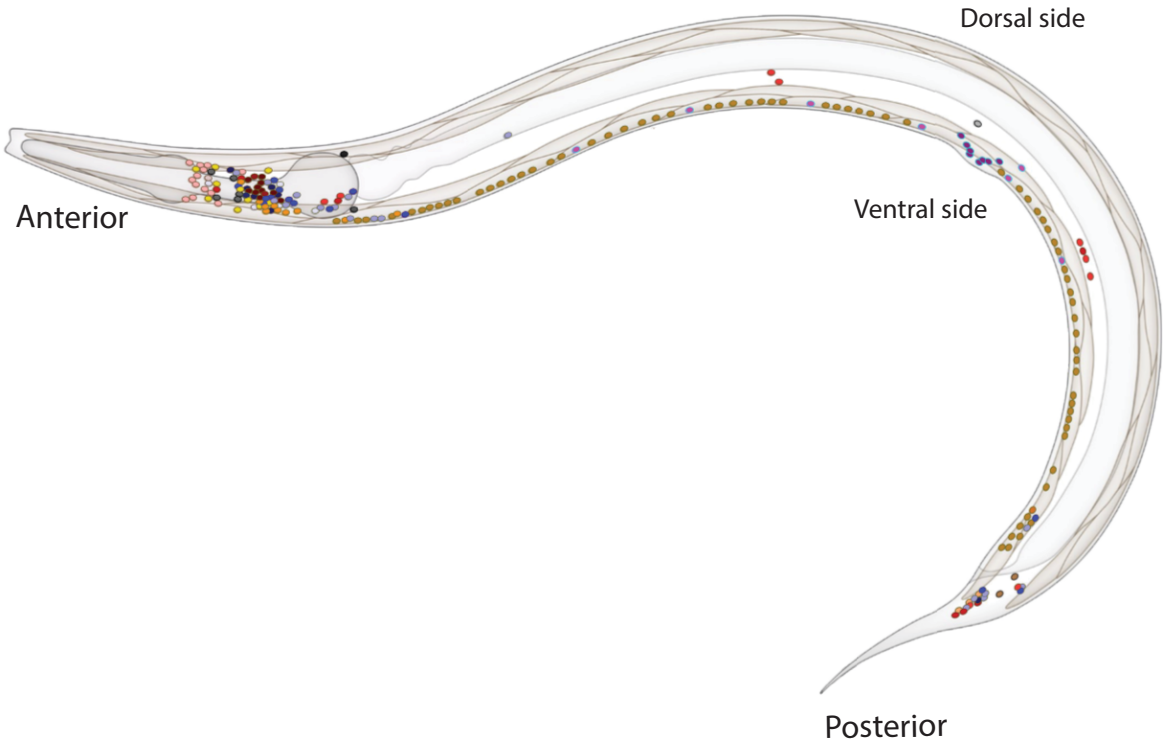

Figure 1.5. *C. elegans* neuronal network in the adult hermaphrodite. Lateral view of the neuronal network in *C. elegans*. Colour cells represent the different neuronal types – Different red shades represent sensory neurons. Different blue shades represent interneurons. Motor neurons are represented in yellow and orange. Finally, purple and pink represent hermaphrodite specific neurons. Adapted from Cook *et al.* (2019). Adapted by permission from Springer Nature Customer Service Centre GmbH from Springer Nature, Nature, Vol.571/issue 7763, Steven J. Cook *et al.* Whole-animal connectomes of both Caenorhabditis elegans sexes; Pages No 63-71, Copyright (2019).

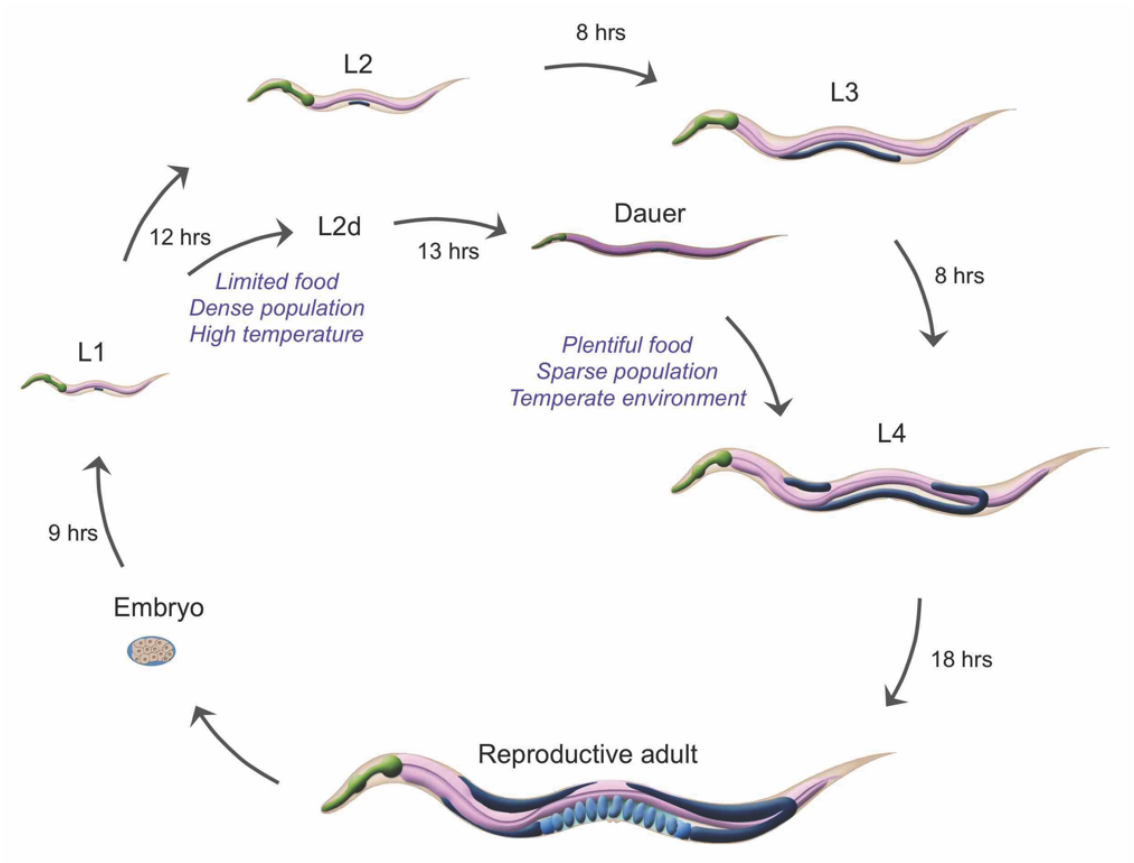

Figure 1.6. *C. elegans* life cycle with development time. This image represents the eggs stage, the 4 larval stages (L1 to L4), the dauer resistant form and the adult hermaphrodite worm. Adapted from Altun *et al.* (2002-2020).

Despite its small size and number of neurons, when compared with the human brain that contains around 86 billion neurons (Braitenberg and Atwood, 1958), approximately 65% of human disease genes have homologues in *C. elegans* (Sonnhammer and Durbin, 1997). Therefore, the *C. elegans* nervous system is an appropriate *in vivo* model for studying neuronal development mechanisms (Sonnhammer and Durbin, 1997). Moreover, *C. elegans* has been extensively used to study and decode the TFs networks required for brain formation for each of the different steps of neuronal development, such as neuronal generation, specification, maintenance and repair (Bülow *et al.*, 2004; Doitsidou *et al.*, 2013; Flames and Hobert, 2009). As an example, *hlh-4* and *egl-13* mutant models were shown to have impaired terminal differentiation of ADL nociceptive neurons and BAG and URX sensory

neurons (Gramstrup Petersen *et al.*, 2013; Masoudi *et al.*, 2018). Further studies revealed that *C. elegans unc-86* and the mouse orthologue Brn3a are essential for the maintenance of the terminal identity of related neurons (Serrano-Saiz *et al.*, 2018). Studies performed by Kratsios *et al.* (2015) showed that the UNC-3 TF is essential to control *madd-4* transcription that is required for neuromuscular junction function in SAB motor neurons. Mutations in *unc-3* results in reduced ACh clusters in the SAB neuronal synapses and to muscle innervation defects (Kratsios *et al.*, 2015).

In this thesis, I focused on a pair of neurons (PVQ left and PVQ right – PVQL/R) that are commonly used as a model to study axon guidance, due to their long axonal process and conserved developmental mechanisms. The PVQs arise during early stages of embryogenesis, approximately 310 minutes after the first cell division, derived from the Abplpa progenitor cell in the posterior end of the embryo (Figure 1.7a) (Sulston *et al.*, 1983; White *et al.*, 1976). To date, very little is known about PVQ fate acquisition. Previous studies performed by Cameron *et al.* (2002) showed that PAG-3 (zn-finger TF) is required for neuroblast and P-cell lineage differentiation. Specifically, this study showed that mutations in *pag-3* lead to dysregulation of PVQ-specific genes, such as sra-6 (Cameron *et al.*, 2002). Moreover, other studies have also identified that ZAG-1 (zn-finger TF) is involved in *sra-6* expression in the PVQs (Cameron *et al.*, 2002; Clark and Chiu, 2003). Similarly, the authors suggest that zag-1 TF is implicated in PVQ specification through PVQ specific gene regulation. However, to date PVQ specification remains poorly understood.

Next, around minute 400 after first cleavage, PVQs start to extend to enter the most posterior extremity of the VNC via the lumbar commissure (Figure 1.7a) (White *et al.*, 1976). Here, the PVQs play an important role in being one of the first neurons to navigate the lumbar commissure, allowing follower axons to navigate the VNC (Durbin, 1987). In the VNC, the PVQs navigate by following the PVP pioneer neurons before terminating in the head by mid-embryogenesis (Figure 1.7a, b & c) (Durbin, 1987).

The close relationship between the PVQs and PVPs was elucidated by studies that revealed that if the PVPR is ablated or misguided it causes misguidance in the PVQL – instead of extending on the left fascicle of the VNC the PVQL joins the PVQR on the right fascicle (Durbin, 1987; Hutter, 2003; Hutter *et al.*, 2005). Interestingly, the opposite effect does not occur, which suggests that if the PVQL is ablated the PVPR still extends correctly and unaffected to its target position (Durbin, 1987).

Despite playing a role in pioneering the lumbar commissure and being a guide for other neurons, how the PVQs are specified is not understood.

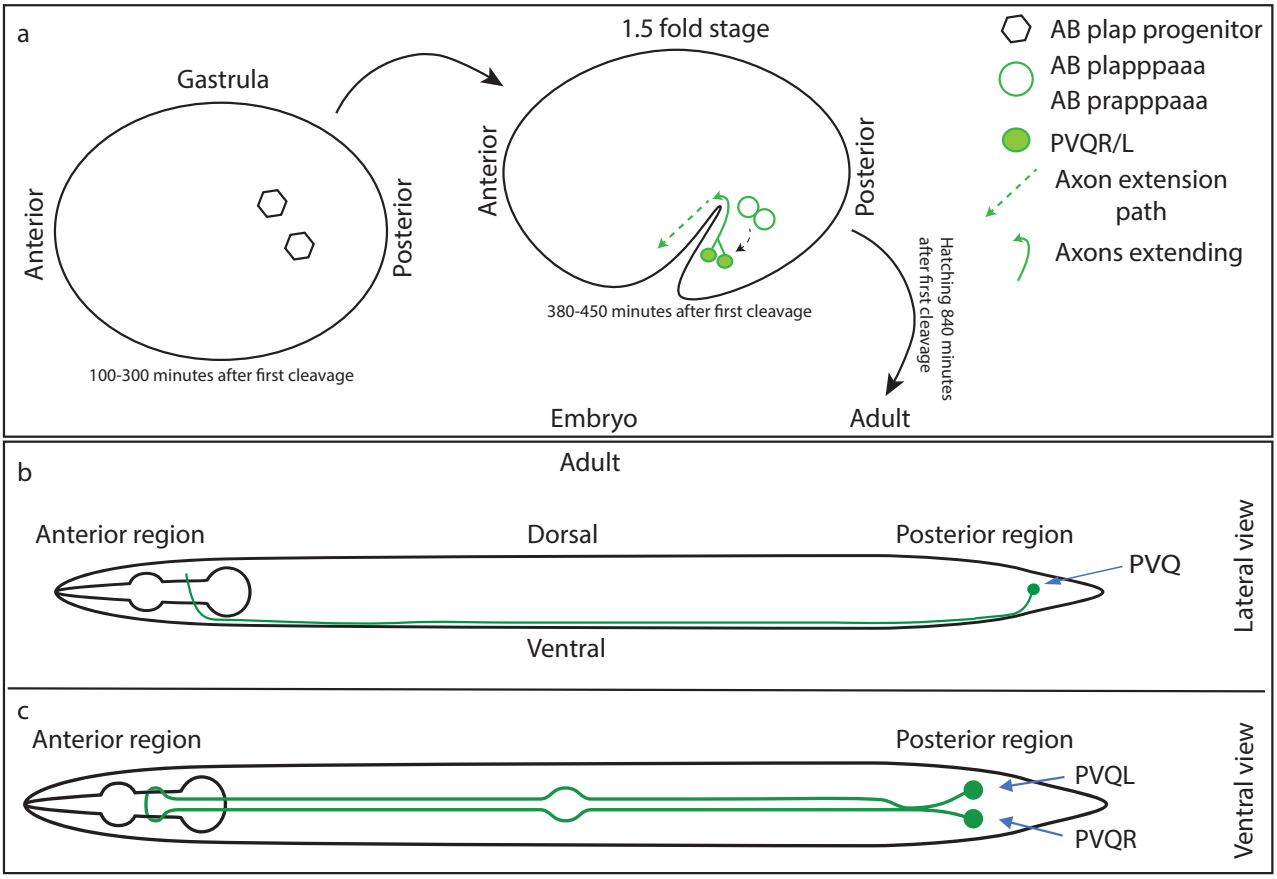

Figure 1.7. Development and projection pattern of the PVQ interneurons in embryo and adult hermaphrodite *C. elegans*. A – PVQ development during embryogenesis. The PVQs are born in the posterior side of the embryo at around 300 minutes after first cleavage. Around minute 400 after first cleavage, PVQs are located in the most posterior region of the embryo. Here, they use the axons of pioneer neurons that are present in the VNC as a guide to project towards the head of the worm. **B –** Lateral view of the PVQs in a young adult. **C –** Ventral view of the PVQs in a young adult. The PVQs are located in the tail of the animal one on the right and another on the left side of the worm.

1.4. Objectives and hypothesis

My PhD project aimed to identify molecular mechanisms that drive brain development and, more specifically, that drive neuronal specification and axon outgrowth. Both of these steps are crucial for nervous system development and function. It is well known that incorrect neuronal specification can lead to uncontrolled cell proliferation that will result in tumour formation. Some examples of disease linked with neuronal specification are gliomas, neuroblastoma and eye disorders. It has been previously shown that dysregulation of TFs activity, such as *ASCL1*, it is one of the underlining causes of gliomas (one of the most common brain tumours) and neuroblastoma (Prasov *et al.*, 2012). On the other hand, mutations in the *ATOH7* TF have a strong link to eye disorders, such as hyperplasia of primary vitreous and bilateral optic nerve aplasia (Prasov *et al.*, 2012). Another crucial step in nervous system neuronal network formation is axon guidance. Axon outgrowth failure is associated with several neurological disorders, including autism spectrum disorder and dyslexia (Anitha *et al.*, 2008; Hannula-Jouppi *et al.*, 2005). Thus, understanding how axons integrate extracellular environmental cues and which molecules are important for their correct migration, outgrowth, and maintenance, will help us to better understand how the nervous system develops and is maintained during adult life.

I aimed to characterize the *sra-6* gene promoter, which drives PVQ expression, to develop a new PVQ reporter for simultaneous examination of multiple PVQ neuron development phases (specification, axon outgrowth and axon guidance). Additionally, this analysis aimed to identify regulatory elements present in the *sra-6* promoter region that are responsible for controlling gene expression in the PVQs. In this work, I identified multiple candidate transcription factors that may be involved in PVQ axon guidance and outgrowth processes. Also, I confirmed that one of the candidate Transcription factors (NFY complex) is involved in PVQ axon guidance and is necessary for the correct generation, specification, or maintenance of the PVQ neurons. Next, I characterized NFY complex function during PVQ development to understand the transcriptional regulation network required for the correct PVQ

development. Finally, I designed and validated CRISPR-Cas9 generated epitope-tagged strains for NFYA-1, NFYB-1 and NFYC-1 to enable proteomics and genomic analysis. This analysis aimed to identify which genes are regulated by the NFY complex and, for the first time, the NFY complex's interactome network in a multicellular organism.

1.4.1. Hypotheses

The molecular networks responsible for neuronal development are vastly uncharacterized. Based on previous work carried in the lab, I hypothesized that the NFY complex plays a crucial role in the development of a pair of glutamatergic interneurons (PVQs). Further, I hypothesized that the NFY complex regulates the battery of genes required to achieve the molecular signature of the PVQs possibly by interacting with specific co-activators/repressors of gene expression specific to the PVQs.

1.4.2. Specific Aims:

1. **Characterization of the *sra-6* gene and development of a genetic tool to investigate the development of PVQ neurons:** To this end, I characterized the promoter for the *sra-6* gene, which is expressed in PVQ, ASH, and ASI neurons. This characterization allowed me to identify potential regulatory regions responsible for controlling PVQ expression during development.

2. **Characterization the role of NFY transcription factor in PVQ development and mutagenesis analysis of promoters for specific putative NFY target genes:** Here, I analysed mutant worms for the *nfya-1*, *nfyb-1,* and *nfyc-1* genes to better understand the role of the NFY transcription factors in PVQ development. This characterization allowed me to better understand the different roles of each of the NFY subunits during PVQ development. Thus, I analysed PVQ specification and axon guidance/outgrowth. Also, I performed localization studies in order to understand if the NFY complex is expressed in

the nuclei of the PVQs. Finally, I used site directed mutagenesis to confirm that specific genes were being directly regulated by the NFY complex.

3. **Proteomics and ChIP-Seq analysis to examine NFY complex function in controlling neuronal fate:** I used CRISPR/Cas9 technology to endogenously tag the NFY complex subunits in order to perform proteomic and ChIP-PCR/ChIP-seq.

Chapter 2. Materials & Methods

2.1. *C. elegans* husbandry

All strains were maintained at 20°C on nematode growth medium (NGM) seeded with 400 µl of *Escherichia coli* (OP50) bacteria, and utilised for experiments after two generations, as previously described (Brenner, 1974).

2.2. *C. elegans* strains

The description of the strains and respective genes used in this study can be found at the Wormbase (wormbase.org) and at the Caenorhabditis Genetics Center (cgc.umn.edu) databases.

Transgenic strains provided by other laboratories are indicated in Table 2.1.

2.3. Microinjection

The transgenic lines used in this study were generated by injection of plasmid DNA, PCR products derived from plasmid DNA, or CRISPR-Cas9 methodology. Injections were performed directly into both hermaphrodite gonads at concentrations described in the respective results chapters. After injections, the worms were kept at 25°C for one hour and then moved to a 20°C incubator. The worms were screened by transgene expression (GFP, RFP or YFP) or PCR as described in section 3.3 and 4.3. Mutant strains and integrated transgenic strains were outcrossed at least three to ten times before being used for further experiments.

The experiments using transgene expression involved the generation of two or three independent transgenic lines to confirm the results. The rescue experiments performed in chapter 4 were done using fosmid DNA constructs available in the Pocock laboratory.

Table 2.1. List of transgenic strains provided by other laboratories used in this work.

Transgenic Strains	Description	Reference
hdIs26	*[Podr-2::GFP; Psra-6::dsRed2]*	(Hutter, 2003)
oyIs14	*[Psra-6:: GFP]*	(Aurelio *et al.*, 2002)
kyIs321	*[Pnpr-11::RFP]*	N/A
stIs11350	*[Pnfya-1::H1-wCherry]*	N/A
stIs10879	*[Pnfyb-1::H1-wCherry]*	N/A
stIs11751	*[Pnfyc-1::H1-wCherry]*	N/A
mjIs27	*[Pmir-124::GFP]*	(Clark *et al.*, 2010)
otEx6860	*[Psrv-32::GFP]*	(Vidal *et al.*, 2018)
otEx6666	*[Psrh-277::GFP]*	(Vidal *et al.*, 2018)
otEx6403	*[Psri-1::GFP]*	(Vidal *et al.*, 2018)
juIs14	*[Pacr-2::GFP]*	(Hallam *et al.*, 2000)
otIs11	*[Pzig-5::GFP + rol-6(su1006)]*	(Aurelio *et al.*, 2003)
otEx233	*[Pdop-1::GFP]*	(Tsalik *et al.*, 2003)
jvsEx428	*[Pegl47A::GFP]*	(Moresco and Koelle, 2004)
rpEx2016	*[Pegl-47B::GFP; Pnpr-11::RFP; Pinx-6::RFP]*	(Moresco and Koelle, 2004)
otIs287	*otIs287 [rab-3(prom1)::2xNLS::YFP + rol-6(su1006)] IV*	(Stefanakis *et al.*, 2015)
otEx6620	*[srg-32prom::gfp, pha-1(+)]*	(Vidal *et al.*, 2017)
otIs339	*[ceh-43(+)(fosmid)::GFP + ttx-3::DsRed + rol-6(su1006)]*	(Doitsidou *et al.*, 2013)
otIs619	*[unc-11(prom8)::2xNLS::TagRFP]*	(Stefanakis *et al.*, 2015)
otIs350	*otIs350 [ric-4(fosmid)::SL2::NLS::YFP::H2B + pha-1(+)]*	(Stefanakis *et al.*, 2015)
otIs619	*unc-11(prom8)::2xNLS::TagRFP] X*	(Stefanakis *et al.*, 2015)

2.4. Brood size analysis

For the brood size experiments, the plates were acclimatized at room temperature and seeded with 50 µl of OP50 one day before use. L4 worms were picked into individuals plates and maintained at 20°C, then transferred into new plates every 24 hours. This process was repeated until the worms ceased to lay eggs, approximately after six to seven days. The total number of worms, viable embryos, and non-fertilized eggs were recorded daily.

2.5. Molecular biology techniques

2.5.1. DNA extraction

DNA extraction was carried out from non-starved plates seeded with OP50. All plates were washed with 400 µl lysis buffer [50 mM KCl, 10 mM Tris pH 8.0, 2.5 mM MgCl2, 0.45 % IGEPAL, 0.45 % Tween-20, 0.01 % gelatin and 0.5 mg/ml Proteinase K (Sigma)], and the samples were incubated overnight at -80°C. On the next day, the samples were incubated at 90°C for 65 minutes, at 95°C for 20 minutes, and stored at -20°C. For DNA extraction from individual worms, a 3-day old worm was picked from the plate and incubated with 16 µl of lysis buffer at -80°C.

2.5.2 Genotyping using PCR

Genotyping was performed using PCR with a set of primers designed to amplify the region of interest (Table 2.2). DNA amplification was performed using an Eppendorf Thermal Cycler using a standard 35-cycle protocol with a denaturation temperature of 95°C, optimized annealing temperature, and an elongation temperature of 72°C. A single PCR reaction (25 µl) consisted of 4 µl DNA (60 ng/µl concentration), 5 µl Betaine, 2.5 µl 10X Standard Taq polymerase Buffer, 0.4 dNTPs (10 mM), 0.5 µl Taq Polymerase, 2 µl primer forward, 2 µl primer reverse, and 9 µl Milli-Q water. The PCR products were visualised on a 1 % agarose gel.

			size		Annealing
Genotyping	**Foward**	**Reverse**	**WT**	**Mutant**	temperature
nfya-1 (ok1174)	cttacaactggaggttggtg	ggatactacacccaaagaac	427 bp	0 bp	58c
	cgtaggtagacatgggttgc	ggatactacacccaaagaac	1961 bp	124 bp	58c
nfya-2 (tm4194)	attggttgaggctgagaggc	acgggcagactcaagctgtg	589 bp	0 bp	58c
	attggttgaggctgagaggc	ttgggacagtgaggtgaatg	825 bp	473 bp	58c
nfyb-1 (cu13)	aggggtaatttgaggcgctg	tcagcgcctcaaattacccc	955 bp	0 bp	58c
nfyb-1 (tm4257)	cggagtttcccggttttacc	tcacatgggtctgctaaatc	592 bp	0 bp	58c
	cggagtttcccggttttacc	cccaaaatggatccaaaacc	970 bp	615 bp	58c
nfyc-1 (tm4541)	cgttcatcctctacgatggc	tttgccggtggtatcgcatc	654 bp	0 bp	58c
	cggtgtacgtgcacacaatg	tgtggtgagatcggttgactg	1007 bp	590 bp	58c
nfyc-1 (rp152)	cagttgcgatgctacagcac	gcttatcacttagtcacctctgctc	0 bp	975 bp	58c
NFYA-1::V5	ccacttcttggtctcgatag	gtaacatgggacgatctatg	0 bp	690 bp	58c
NFYB-1::HA	ccatacgatgttccagattacg	catgattcggaccacgttagc	0 bp	824 bp	58c
NFYC-1::HA	ccatacgatgttccagattacg	tcactctagccatgggaacg	0 bp	474 bp	58c
NFYC-1::Stopcassette	cagttgcgatgctacagcac	gcttatcacttagtcacctctgctc	0 bp	975 bp	58c

Table 2.2. Primers and expected product sizes for the different mutant strains.

2.5.3. RNA extraction

RNA extraction was carried out using the RNeasy kit (Qiagen), according to the manufacturer's instructions. Briefly, each N2 plate was washed three times with 1 ml of M9 solution (22 mM KH2PO4, 50 mM Na2HPO4, 86 mM NaCl, 1 mM MgSO4) and 1.5 ml of LS Trizol was added per 500 µl of pellet. The samples were immediately frozen in liquid nitrogen, thawed at 37°C, and subsequently homogenised using a vortex mixer (repeated seven times). 200 µl of chloroform was added per 1 ml of homogenate and the sample was vortexed for 15 seconds. The samples were incubated at room temperature for 5-15 minutes and centrifuged in 1200 g at 4°C for 30 minutes. The aqueous phase was transferred to a new tube and the RNA was washed by adding 100 % ethanol in a 1.5:1 ratio (ethanol: homogenate) to each sample. The homogenate was transferred to a spin column and the reaction was centrifuged in 17900 g at 4°C for 15 seconds (flow-through was discarded). 700 µl RW1 buffer was added to the columns and the samples were centrifuged during 15 seconds at ~17,900 g (flow-through was discarded). 80 µl of DNase+RDD buffer was added directly into the centre of the membrane and left at room temperature for 15 minutes. 500 µl of RWL buffer was added to the samples and centrifuged for 15 seconds at ~17,900 g (the flow-through was discarded). 500 µl of RPE buffer was added to the samples and centrifuged for 2 minutes at ~17,900 g (the flow-through was discarded). The samples were centrifuged for 2 minutes at ~17,900 g. The column was transferred to a new collection tube and 30 µl of RNase-free water was added to the centre of the membrane followed by a centrifugation step at ~17,900 g for 1 minute. The sample RNA concentration was measured using the NanoDrop1000 Spectrophotometer (Thermo Fisher Scientific).

2.5.4. cDNA synthesis

cDNA synthesis was performed using the ImProm II synthesis kit (Promega) according to the manufacturer's guidelines. A 5 µl reaction containing 500 ng of template RNA, 1 µl of oligonucleotide/random primer mix (1:3 ratio; 0.5 µg/ µl) and RNase free water was incubated for 5 minutes at 70°C, followed by a cooling step in ice for 5 minutes 15 µl of a mix containing 6.5 µl RNA free water, 4 µl of 5X ImProm II buffer, 1 µl dNTPs

(10mM), 0.5 µl RNasin ribonuclease inhibitor, 2 µl of $MgCl_2$ (25 mM), and 1 µl of ImProm II enzyme was added and the samples were then incubated at 25°C for 5 minutes, at 42°C for 1 hour, followed by inactivation of the reverse transcriptase enzyme at 70°C for 15 minutes.

2.5.5. DNA PCR amplification and purification (promoter or cDNA amplification)

For the development of the reporters and rescue constructs, it was necessary to amplify the promoter regions of the genes from wild type DNA or amplify the coding regions from cDNA prior to cloning. The amplification was performed using the Phusion High-Fidelity DNA Polymerase (New England Biolabs), according to the manufacturer's instructions. The primers used in this work were designed using the serial cloner (SerialBasics, version 2.6.1), the In-Fusion Cloning tool (www.takarabio.com/learning-centers/cloning/in-fusion-cloning-tools) and the Oligocalc online tool v3.27 (biotools.nubic.northwestern.edu). A 50 µl reaction containing 31 µl of ultrapure water, 10 µl of 5X Phusion HF Buffer, 3 µl of each primer, 1 µl of dNTPs (10 mM), 1 µl of Phusion DNA Polymerase, and 1 µl of template DNA/cDNA (60-150 ng). DNA amplification was performed using an Eppendorf Thermal Cycler using a standard 35-cycle protocol with a denaturation temperature of 95°C, annealing temperature of 56-60°C, and an extension temperature of 72°C. The PCR products were separated on a 1 % agarose gel.

The bands were excised from the gel and the DNA was purified with a QIAquick DNA purification kit (Qiagen) accordingly to the manufacturer protocol. 1 ml of Buffer QG was added and the samples were incubated at 65°C for 15 minutes in a thermomixer with vigorous shaking, followed by incubation at 4°C for 4 minutes. The samples were transferred to QIAquick spin columns and centrifuged at ~10,600 g during 60 seconds and the flow-through was discarded. The QIAquick spin columns were washed by adding 750 µl Buffer PE and a centrifugation step at ~10,600 g for 60 seconds. The samples were washed a second time by adding 750 µl of PE Buffer to the columns, incubated for 5 minutes at room temperature, and centrifuged at ~10,600 g for 60

seconds (the flow-through was discarded). The samples were centrifuged at ~17,900 g for 1 minute to remove all residues of PE buffer. Finally, the QIAquick spin columns were placed in a clean 1.5 ml collection tube. The samples were eluted by adding 30 to 50 µl of Buffer EB (10 mM Tris-HCl, pH 8.5), followed by a centrifugation step at ~17,900 g for 2 minutes.

2.5.6. Cloning

The amplified inserts and plasmid backbones (for example, pPD95.75 or pPD49.26) were digested using specific restriction enzymes, shown in Table 2.3. A single restriction enzyme reaction (50 µl) consisted of 600 ng of DNA, 1 µl of Fast Digest (Thermo scientific) restriction enzyme, 2 µl of 10X FastDigest Buffer/FastDigest Green Buffer, and MilliQ water. The plasmid backbone was additionally treated with FastAP Thermosensitive Alkaline Phosphatase after digestion to avoid re-ligation and increase cloning efficiency.

Cloning of inserts was performed using the Quick protocol or restriction free protocol from Takara-bio Inc. For Quick protocol, the ligation reaction (10 µl) was set up on ice by adding 1 µl of 2X Rapid Ligation Buffer, 1 µl of T4 DNA ligase, and a 7:1 ratio of insert and plasmid. The reaction was then incubated overnight at 15°C.

For the restriction free cloning protocol, the insert was amplified using primers (Table 2.3) designed with the In-Fusion Cloning Tool (Takara-bio Inc.) as described in section 2.5.5. Subsequently, 5 µl of the PCR mixture was treated with 2 µl cloning enhancer (Takara-bio Inc.). The cloning reaction was setup by adding 2 µl of 5X Ib-Fusion Premix, 1 µl linearized vector, 2 µl insert (treated with cloning enhancer), and H_2O in a total volume of 10 µl. The reaction was incubated for 15 minutes at 50°C.

Table 2.3. Primers and restriction enzymes used for the generation of the plasmids.

Number	Lab	Name	Resistance	Enzymes	Forward	Reverse
RJP703	Pocock Lab	*OPM2.5::WrmScarlet (49.26)*	Amp	Enzymes - OPM2.5 (XbaI; SmaI);	OPM1F - aaaaTCTAGAtgtaagacatacgttagctgg	OPM2R - aaaaccc gggGCATCGGCCGAGTAAAACTT
				WrmScarlet - (NheI; KpnI)	WrmScarlet F - aaaagctagcATGGTCAGCAAGGGAGAGGC	WrmScarlet R - aaaaggtaccATTACTTGTAGAGCTCGTCC
RJP704	Pocock Lab	*Pnpr-11::WrmScarlet (49.26)*	Amp	Enzymes - Pnpr-11 (XbaI; SmaI);	Pnpr-11 F_XbaI - AAAATCTAGAGTACGGCTTCTGTGGGAGGC	Pnpr11 R _SmaI - AAAACCCGGGGATCCCATTTGTTGAAAGTTC
				WrmScarlet - (NheI; KpnI)	N/A	N/A
RJP705	Pocock Lab	*OPM2.5::GFP (pPD117.01)*	Amp	Enzymes - OPM2.5 (XbaI; SmaI)	OPM1F - aaaaTCTAGAtgtaagacatacgttagctgg	OPM2R - aaaaccc gggGCATCGGCCGAGTAAAACTT
RJP706	Pocock Lab	*Psra-6::GFP (pPD117.01)*	Amp	Enzymes - Psra-6 (XbaI; SmaI)	N/A	N/A
RJP707	Pocock Lab	*OPM2.5::NeonGreen (49.26)*	Amp	Enzymes - OPM2.5 (XbaI; SmaI); Neongreen - (NheI; KpnI)	Neongreen F NheI - aaaaggtaccCTACTTGTAGAGTTCATCC	Neongreen R KpnI - aaaaggtaccCTACTTGTAGAGTTCATCC
RJP708	Pocock Lab	*Pnpr-11::nfya-1 cDNA(49.26)*	Amp	Enzymes - Pnpr-11 (XbaI; SmaI);	N/A	N/A
				nfya-1- (KpnI; NcoI)	OPM64F_nfya_KpnI_nfya_cDNA - aaaaggtaccATGAATGGAGCGTCGAGGGG	OPM65R_nfya_NcoI - aaaaccatggTTAGAGATTCGTGAAACTTTG
RJP709	Pocock Lab	*Pnpr-11::nfyb-1 cDNA(49.26)*	Amp	Enzymes - Pnpr-11 (XbaI; SmaI);	N/A	N/A
				nfyb-1- (NheI; KpnI)	OPM60F_nfyb_cDNA - aaaagctagcATGGATCCAAAACCCATAAAC	OPM61R_nfyb_cDNA - aaaaggtaccCTACTTTTTCCTCTTCGTCGGC
RJP710	Pocock Lab	*Pnpr-11::nfyc-1 cDNA(49.26)*	Amp	Enzymes - Pnpr-11 (XbaI; SmaI);	N/A	N/A
				nfyc-1- (NheI; KpnI)	OPM62F_nfyc_cDNA - aaaagctagcATGTCGCAATTCCCAGAAG	OPM63R_nfyc_cDNA - aaaaggtaccTCAATTGTGGTGAGATCGG
RJP711	Pocock Lab	*Phsp-16.2::nfya-1 cDNA(49.78)*	Amp	nfya-1- (KpnI; NcoI)	N/A	N/A
RJP712	Pocock Lab	*Punc-86::nfya-1 cDNA (49.26)*	Amp	nfya-1- (KpnI; NcoI)	N/A	N/A
RJP713	Pocock Lab	*Pnpr-11-1-RPL-22-HA (pSF214)*	Amp	Enzymes - Pnpr-11 (XbaI; SmaI);	N/A	N/A
N/A	Pocock Lab	*Psra-6_sid-1 cDNA(pD49_26)*	Amp	Enzymes - Psra-6 (XbaI; SmaI);	N/A	N/A
				sid-1 - (NheI; KpnI)	OPM24F - aaaagctagcATGATTCGTGTTTATTTGATAATT	OPM25R - aaaaccatggCTAGAAAATGTTAATCGAAGTT
RJP714	Pocock Lab	*Psri-1::GFP (pPD95.75) sample 1*	Amp	Enzymes - Psri-1 (XbaI; SmaI)	OPM94F - aaaaTCTAGAaaattgcattatattaatgttgttc	OPM94R - aaaaCCCGGGTTTTTAGTATGCTAAACTCGC
RJP715	Pocock Lab	*Psri-1::GFP (pPD95.75) (Mutagenesis on the CCAAT site). Used sample 2*	Amp	Psri-1 Mutagenesis	OPM95F - Psri -1 CCCTCGCTGATGAAATACTGCTTTGACATTTTTGC	OPM96R - Psri - 1TTTCATCAGCGAGGGAAATTGAATTTTAATTCG
RJP716	Pocock Lab	*Psrh-277::GFP (pPD95.75). Samle 2*	Amp	Enzymes - Psrh-177 (XbaI; SmaI)	OPM97F - aaaaTCTAGAgttgcacagccggttatgcg	OPM98R - aaaaCCCGGGttgtggcgacggcgacca
RJP717	Pocock Lab	*Psrh-277::GFP (pPD95.75) (Mutagenesis on the CCAAT site). Used sample 1*	Amp	Psrh-277 Mutagenesis	OPM101F - Psrh-277 TTACACCTGCTCCTAGAAAACCTACAATTTGTGA	OPM102R - Psrh-277 TAGGAGCAGGTGTAATTATGTTTCACGAGGGCT
RJP718	Pocock Lab	*Psrg-32::GFP (pPD95.75) sample 2*	Amp	RF cloning Enzymes - (XbaI; SmaI)	srg-32 F (RF) - GCAGGTCGACTCTAGACAATTTATCGCACTGATCT CTGAAG	PM srg-32 R (RF) - CCTTTGGCCAATCCCGGGGTTGAGGGGCGCGAA GAG

2.5.7. Transformation of competent cells using heat shock

Frozen aliquots of competent cells (DH5α) were thawed on ice for 5 minutes. 1 µl of cloning or re-transformation sample (1 pg to 100 ng) was incubated with 40 µl of competent cells on ice for 30 minutes. A heat shock was applied at 42°C for 30 seconds, followed by cooling on ice for 5 minutes. Then, 400 µl of SOCs outgrowth or LB medium was added and the bacterial culture was incubated for 1 hour with agitation at 37°C. Bacteria were plated on solid medium (LB) containing specific antibiotic solution (100 µg/µl), and grown overnight in an incubator at 37°C.

2.5.8. Plasmid DNA Isolation

Fosmid and plasmid samples were prepared based on the FosmidMAXTM and QIAprep Miniprep DNA Purification protocol. A single colony was collected from the selective plate and inoculated in a 15 ml Falcon tube containing 1-5 ml LB medium supplemented with the appropriate antibiotic concentration (chloramphenicol, ampicillin or kanamycin). The samples were then incubated overnight with agitation at 37°C.

The bacterial cells were harvested by centrifugation at 6,800 g for 1 minute, and the supernatant was discarded. Next, 250 µl of the Buffer P1 (RNase - 100 µg/ml and LyseBlue) was mixed with the pellet, and 250 µl of the Buffer P2 was added. Each sample was mixed by inverting the tube 4-6 times. 350 µl of the Buffer N3 was added and the samples were again mixed by inverting the tubes 6-8 times, followed by centrifugation at 17,900 g for 10 minutes. After these steps, the supernatant was transferred to a QIAprep 2.0 spin column for further centrifugation for 30-60 seconds, and the flow-through was discarded. The QIAprep 2.0 spin column was washed with 500 µl PB buffer, followed by centrifugation at 10,600 g for 60 seconds, and 750 µl of PE buffer, and another centrifugation at 10,600 g for 60 seconds. A second wash with PE buffer was performed by incubation at room temperature for 5 minutes, followed by centrifugation at 10,600 g for 60 seconds. The ethanol contained in both washing solutions were allowed to evaporate by centrifugation at 17,900 g for 1 minute. Finally, the DNA was eluted by placing the QIAprep 2.0 column in a new collection tube,

adding 30-50 µl of the EB buffer (10 mM Tris-HCl, pH 8.5), and centrifuging for 2 minutes at 17,900 g.

The DNA concentration was quantified using the NanoDrop1000 Spectrophotometer (Thermo Fisher Scientific) and stored at -20°C for further use. The purified plasmids and fosmids were sent for sanger sequencing for confirmation of its sequence (Table 2.4).

Table 2.4. Primers used for sequencing the different plasmids generated.

Sequencing primers	Sequence	
pPD49.26 F	caaaggacccaaaggtatgtttcg	49.26 Sequencing
pPD49.26 R	agagtaattggacttagaagtcagagg	
pPD95.75 F	atgaccatgattacgccaagc	95.75 Sequencing
pPD95.75 R	ttccgtatgttgcatcacc	
OPM54F_Pnpr-11_seq	ggctgaaatcactcacaacg	Pnpr-11
OPM55F_Pnpr-11_seq	cgaaaaacatgaaatggtgg	
OPM56F_Pnpr-11_seq	ccttatttcttcaagttggg	
OPM57F_Pnpr-11_seq	caatggagatgtggtcagtg	
OPM58F_nfya_cDNA	aaaagctagcatgaatggagcgtcgagggg	NFYA-1_cDNA
OPM59R_nfya_cDNA	aaaaggtaccttagagattcgtgaaactttg	
OPM87F_nfyb_HA_Seq	cgatttagcagaccaatgtg	NFYB-1 HA_CRISPR
NFYB_HA_RseqNew1	cctccagtaacattccttcg	
NFYB_HA_RseqNew2	cggatgatgaagctccattg	
OPM91F_nfyb_HA_Seq	gcaattttcagctcatttcc	
OPM88F_nfyc_HA_Seq	gcctcggtgtacgtgcacac	NFYC-1 HA_CRISPR
OPM89F_nfyc_HA_Seq	ctgacgcctgaaacatttcg	
OPM90F_nfyc_HA_Seq	cagttgcgatgctacagcac	
OPM88F_nfyc_HA_Seq	gcctcggtgtacgtgcacac	NFYC-1::Stopcassette
OPM89F_nfyc_HA_Seq	ctgacgcctgaaacatttcg	
Prab-3R_Seq	cacttctgaggagtgtatgc	Rab-3 CCAAT site seq
ROP01-seq-s1	gtcgttgctactagggttcg	NFYC-1A V5_CRISPR

2.5.9. Site-directed mutagenesis

The plasmids for the site-directed mutagenesis experiments were selected and the site-directed mutagenesis kit (Takara-bio Inc.) was used to alter specific binding sites

on the promoter regions of the genes, according to the manufacturer's protocol. First, a reporter plasmid with the promoter and the fluorescence reporter needed was generated (Table 2.3). Next, the In-Fusion Cloning Tool (Takara-bio Inc..) was used to design primers that would induce the desired mutation (substitution, deletion or insertion) (Table 2.3). A PCR reaction was then prepared as described in section 2.4.4, and 2 µl of cloning enhancer (Takara-bio Inc.) was added to 5 µl of PCR reaction volume to remove any original plasmid used. The samples were incubated at 37°C for 15 minutes, at 80°C for 15 minutes, and stored at -20°C for further use.

A 10 µl In-Fusion Cloning reaction (Takara-bio Inc.) was prepared by adding 2 µl of 5X In-Fusion HD Enzyme Premix, 1-2 µl of Purified PCR fragment and water, followed by an incubation for 15 minutes at 50°C. The samples were then processed for transformation as described in section 2.4.6, the plasmid DNA was isolated and sent for Sanger sequencing.

2.6. CRISPR/Cas9 methodology for endogenous gene manipulation

Transgenic strains were generated using CRISPR-Cas9 genome editing following the protocol outlined by Dokshin *et al.* (2018). The target sites for CRISPR-Cas9 genome editing were identified using the software from Integrated DNA Technologies, Inc. (IDT). The oligonucleotides and repair templates, also known as synthetic single-stranded oligodeoxynucleotide (ssODN), were ordered from Sigma (Table 2.5). Synthesis of the gRNAs was performed using the protocol outlined in Dokshin *et al.* (2018) using injection mix consisting of 0.5 µl of 10 µg/µl Cas9 , 5 µl of 0.4 µg/µl tracrRNA, 2.8 µl of 0.4 µg/µl crRNA, 2.2 µl of 1 µg/µl ssODN (for the NFYC-1 stop cassette)/ 10 µl of 200 ng/µl dsDNA (for FlagTag, HA or GFP tagging), 0.8 µl of *Pmyo-2*::*mCherry* (co-injection marker), and H_2O in a final volume of 20 µl.

Note: The oligonucleotides used in this study are depicted in the table below.

After injection, two F1 plates were selected and 100 to 200 F2 worms were randomly picked to individual plates. DNA was then extracted from worms (section 2.5.1) and the samples were then genotyped by PCR as described in section 2.5.5.

Table 2.5. CRISPR ssODN/dsDNA and sgRNA sequences

CRISPR Strain	ssODN/dsDNA	sgRNA
NFYC-1 - Stop-IN cassette	GGGAAGTTTGTCCAGAGCAGAGGTGACTAAGTGATAAGCTAGC	TTCTGGGAATTGCGACATCA
V5::NFYA-1	GGAAAACCAATTCCAAATCCACTTCTTGGTCTCGATAGTACT	N/A
Flagtag::NFYB-1	CCCGACTACAAAGACGATGACGACAAGGACTACAAAGACGATG ACGACAAGGACTACAAAGACGATGACGACAAG	CCGTCTAAAAGCTCCCAAAA
HA::NFYB-1	TACCCATACGATGTTCCAGATTACGCT	CCGTCTAAAAGCTCCCAAAA
HA::NFYC-1	TACCCATACGATGTTCCAGATTACGCT	TTCTGGGAATTGCGACATCA

2.7. Proteomics

2.7.1. Protein extraction

To obtain protein lysates, the worms were grown in NGM plates for 3-4 days. Then, the plates were washed with M9 buffer to collect the worms, and the samples were centrifuged at 4487 g to remove the M9 buffer and traces of bacteria (repeated 3 times). After this step, 3 ml of cold lysis buffer was added and the samples were again centrifuged at 2655 g. 500 µl of cold lysis buffer was added and the samples were transferred into low protein binding tubes (Eppendorf), followed by sonication for 5 to 10 cycles or mechanical grinding at 4°C. To remove cell debris, the samples were centrifuged at 2655g at 4°C and the supernatant was collected. The samples were frozen with liquid nitrogen and stored at -80°C.

Lysis Buffer solution (make this solution one day before use and keep it at 4°C):

Tris pH 7.4 – 250 µl (50 mM)

NaCl – 750 µl (150 mM)

Triton X-100 – 100 µl (2 %)

SDS – 10 µl (0.02 %)

Protease inhibitor – 250 µl **(added just before use)**

Ultra-pure water – 3640 µl

Total volume – 5 ml

2.7.2. Western Blot

The protein lysates, along with solution 1 containing reducing agent (Thermo Scientific) and protein loading dye (Thermo Scientific), were denatured at 95°C for 10 minutes and subsequently separated by SDS-PAGE in a 10 % acrylamide gel at 70V. The proteins were transferred onto a PVDF membrane (Thermo Scientific) using a iBlot™ 2 Gel Transfer (Thermo Scientific) device for 10 minutes, and the PVDF membrane was incubated in 5-10 ml 5 % BSA for 1 hour at room temperature. The PVDF membrane was then probed with primary antibody (in 1X PBST solution) at 4°C overnight. Primary antibodies used in this study were: Anti-V5 mouse IgG_2a (Bio-Rad, 1:1000), Anti-HA High Affinity rat IgG_1 (Roche, 1:500).

On the following day, the PVDF membrane was washed 5 times for 5 minutes in 1X PBST and incubated with HRP-conjugated goat anti-mouse IgG secondary antibody (Thermo Scientific, 1:1000) for 1 hour at room temperature. The PVDF membrane was washed 5 times for 5 minutes in PBST, and the immunoblot was developed using the Pierce™ ECL Western Blotting Substrate (Thermo Scientific).

Solution 1:

LDS (4X) – 125 µl

Mili-Q water – 100 µl

Reducing agent (10X) – 25 µl

Total volume – 250 µl

PBST 1X:

PBS1X – 1 L

Tween 20 – 1 ml

2.7.3. Immuno-Precipitation (IP)

Before the immunoprecipitation protocol, 20 µl of each sample was collected to be used as input control and stored at -80°C. The remaining protein lysate was treated with 40 µl of magnetic beads (Dynabeads, Thermo Scientific) for 15 minutes to remove unspecific binding of proteins to the beads. The beads were then magnetised with a magnet and the supernatant was transferred into a low protein binding tube (Eppendorf). The supernatant was incubated with 300 µl of ice cold lysis buffer and 1 µg of primary antibody at 4°C overnight, while the beads were stored at -80°C and later used for control of the immunoprecipitation protocol. Next, 20 µl of the magnetic beads was added to the sample, followed by incubation at 4°C for 5 hours or overnight. The magnetic beads were magnetised, and the supernatant was transferred to a new tube and stored at -80°C (control). Wash Buffer 1 was added to the beads for 5 minutes, then wash buffer 2 was added for 5 minutes, and wash buffer 3 was added twice for 5 minutes. Finally, the proteins were eluted from the beads using 0.2 M glycine pH 2.5 and the pH was neutralized by adding 1M Tris-HCl pH 8. The eluted samples and beads were stored at -80°C.

Wash Buffer 1:

Tris pH 7.4 – 250 µl (50 mM)

NaCl – 750 µl (150 mM)

Protease inhibitor – 250 µl (added just before use)

Ultra-pure water – 3750 µl

Total volume – 5 ml

Wash Buffer 2:

Tris pH 7.4 – 250 µl (50 mM)

NaCl – 2250 µl (450 mM)

Protease inhibitor – 250 µl (added just before use)

Ultra-pure water – 2250 µl

Total volume – 5 ml

Wash Buffer 3:

Tris pH 7.4 – 250 µl (50 mM)

Protease inhibitor – 250 µl (added just before use)

Ultra-pure water – 4500 µl

Total volume - 5 ml

2.8. Analyses of GFP expression and neuron morphology

For analyses of GFP expression and neuron morphology, L1 larvae and L4 hermaphrodites were scored. Accordingly, L1 larvae and L4 hermaphrodites were anesthetized in 10 mM sodium azide (NaN_3) and mounted on a 5 % agarose glass slide. Images were capture using an automated fluorescence microscope (Zeiss, AXIO) and the ZEN software.

2.9. PVQ guidance and specification defects

Guidance and outgrowth defects in the PVQs were classified into four categories: guidance defects occurring due to misguidance of the PVQL, guidance defects occurring due to misguidance of the PVQR, single stop for outgrowth defect or double stop for outgrowth defect (PVQL or PVQR). Specification defects were divided into 3 categories: PVQs that present GFP (PVQs ON), PVQs without GFP (PVQs OFF) and PVQS with a weak expression of GFP when compared with controls (PVQs weak).

2.10. Calculation of the Corrected Total Cell Fluorescence (CTCFs)

In order to quantify the amount of fluorescence present in each neuron, L4 animals with target transgenes were selected and mounted for microscopy as described in section 2.2. DIC and fluorescent images of the ventral side (for PVQs) and lateral (for motor neurons) of the animals were captured using a 40X objective. Fluorescence was

then quantified using ImageJ (National Institute of Health, USA, version 1.52h). After the selection of the area of fluorescence and a background region, a measurement was performed. Based on the data, a CTCF calculation was carried out using Excel (Microsoft®) (CTCF = Integrated density – (selected area x mean fluorescence of background readings).

2.11. Statistical analyses

Statistical analyses were performed using the GraphPad Prism 7 software (graphpad.com). The tests used were the t-student test or One-way ANOVA for multi-comparisons, and the values were expressed as mean ± standard error of the mean (SEM) or mean ± standard deviation (SD). Differences with a p value<0.05 were considered significant.

Chapter 3. Characterization of neuron-specific transcriptional elements on the promoter of the *sra-6* gene in *C. elegans*

3.1. Abstract

The establishment of correct brain architecture during development is an exceptionally complex process requiring precisely controlled cell specification, migration, axon outgrowth and axon guidance. These events are controlled by multiple transcription factors and conserved guidance systems. The *Caenorhabditis elegans* nervous system is an excellent model to study brain development due to its relative simplicity and conserved nature of development. My studies investigate how the cell fate of a pair of *C. elegans* glutamatergic interneurons, called PVQ, is controlled.

In this third chapter, I examine potential transcriptional mechanisms that regulate PVQ-specific gene expression. To this end, I characterised cis-regulatory elements of the *sra-6* promoter that are responsible for driving PVQ expression. This work identified conserved motifs within the *sra-6* promoter that are predicted to be bound by specific transcription factors, including the NFY transcription factor complex.

In the next chapter, I will perform in-depth analysis of the NFY complex in order to dissect its function during PVQ development.

3.2. Introduction

Understanding how neuronal differentiation is regulated and the mechanisms that control neuron-specific gene expression is an essential and challenging question in developmental neurobiology. The generation and differentiation of specific neurons, such us motor and interneurons, is controlled by transcription factors (TF) that spatially and temporally regulate gene expression. TFs bind to specific motifs located within promoters, introns and downstream regions of target gene in order to enhance or repress gene expression (Spitz and Furlong, 2012). I am interested in elucidating mechanisms that control the regulatory landscape of neuronal specific genes. As a model, I analyse how the *serpentine receptor class alpha-6* (*sra-6)* gene is transcriptionally regulated such that its expression is driven in three pairs of bilaterally symmetric neurons, ASHL/R, ASIL/R and PVQL/R. *sra-6* encodes a G protein-coupled receptor (GPCR) with an unknown function in *C. elegans*. GPCRs form the largest group of cell surface proteins in humans, and play an essential role in sensory neuronal function. GPCRs are known to bind to different ligands, from odorants to hormones and cytokines, and are major targets for pharmaceutical therapies. In *C. elegans,* GPCRs are involved in behavioural responses to the environment, such as detecting food cues and noxious substances.

In this study, I characterized the promoter for the *sra-6* gene in order to identify mechanisms required for controlling *sra-6* expression during development. This characterization allowed me to identify regulatory regions in the *sra-6* promoter responsible for controlling the expression of *sra-6* gene in the PVQ neurons. Moreover, this analysis allowed me to develop a new neuronal reporter to enable us to simultaneously study multiple phases of PVQ neuron development (specification, axon outgrowth, axon guidance development). So far, I have identified candidate transcription factors that may be involved in the axon guidance and outgrowth processes. Moreover, I confirmed that one of the candidate genes is involved in PVQ axon guidance and is necessary for the correct generation, specification, or

maintenance of the PVQ neurons. Further analysis into this mechanism will be discussed in more detail in chapter 4 and chapter 5.

3.3. Methods

Strains

All Strains were maintained at 20°C for two generations prior to scoring as previously described (Brenner, 1974), unless stated otherwise. The transgenic strains used in this work are detailed in Table 1. Mutant strains used: *cfi-1 (ky651), oyIs14; ceh-43 (ot406), oyIs14;*

sra-6 sequence

The *sra-6* promoter sequence from the *C. elegans* genome was downloaded from WormBase (https://www.wormbase.org/species/c_elegans/). The *Psra-6* region comprises 3775bp upstream of the predicted ATG transcriptional start site. This region was amplified by PCR, cloned into the pPD95.75 and pPD117.01 plasmids and posteriorly confirmed by sequencing. We then use this template promoter sequence to generate smaller fragments, OPM3 to OPM2.93, through PCR amplification that englobes a region of the *sra-6* promoter, full GFP sequence and full *unc-54* or *let-858* 3'UTR. All these fragments were extracted from agarose gel and subsequently purified with a Qiagen purification kit. Posteriorly, all fragments and plasmids were co-injected at 50ng/ul with *pmyo-2::RFP* reporter at 5ng/ul and 100ng/ul of Bacterial DNA into the germ line of WT hermaphrodite worms (Primers available in Supplemental Table 1).

Fluorescence microscopy

GFP expression and neuron morphology was scored in embryos and L4 hermaphrodites by mounting the worms onto a 5% agarose pad on a glass slide and anesthetized them in 10 mM Sodium Azide (NaN_3). Images were capture using an automated fluorescence microscope (Zeiss, AXIO) and ZEN software.

Scoring of crossover defects

The guidance or outgrowth defects of PVQs were classified into the following categories: guidance defect occurring due to misguidance of the PVQL or PVQR, or single or double stop for outgrowth defect (PVQL or PVQR).

3.4. Results

In order to study how the *sra-6* gene is transcriptionally regulated in the PVQ neurons during development, I examined what elements of the *sra-6* promoter (*Psra-6*) are required to drive expression. The starting point of this analysis was to generate a transgenic construct comprising the full-length promoter of *sra-6* driving the expression of GFP using the pPD95.75 plasmid backbone. This construct was used to generate a transgenic line expressing GFP in the ASH, ASI and PVQ neurons (Figure 3.1). In order to uncover the one or more regions responsible for driving expression specifically in the PVQ neurons, I designed primers to generate shorter fragments of the *Psra-6*::GFP by PCR (Figure 3.1). These PCR products were used to generate transgenic animals to enable characterization of the expression pattern.

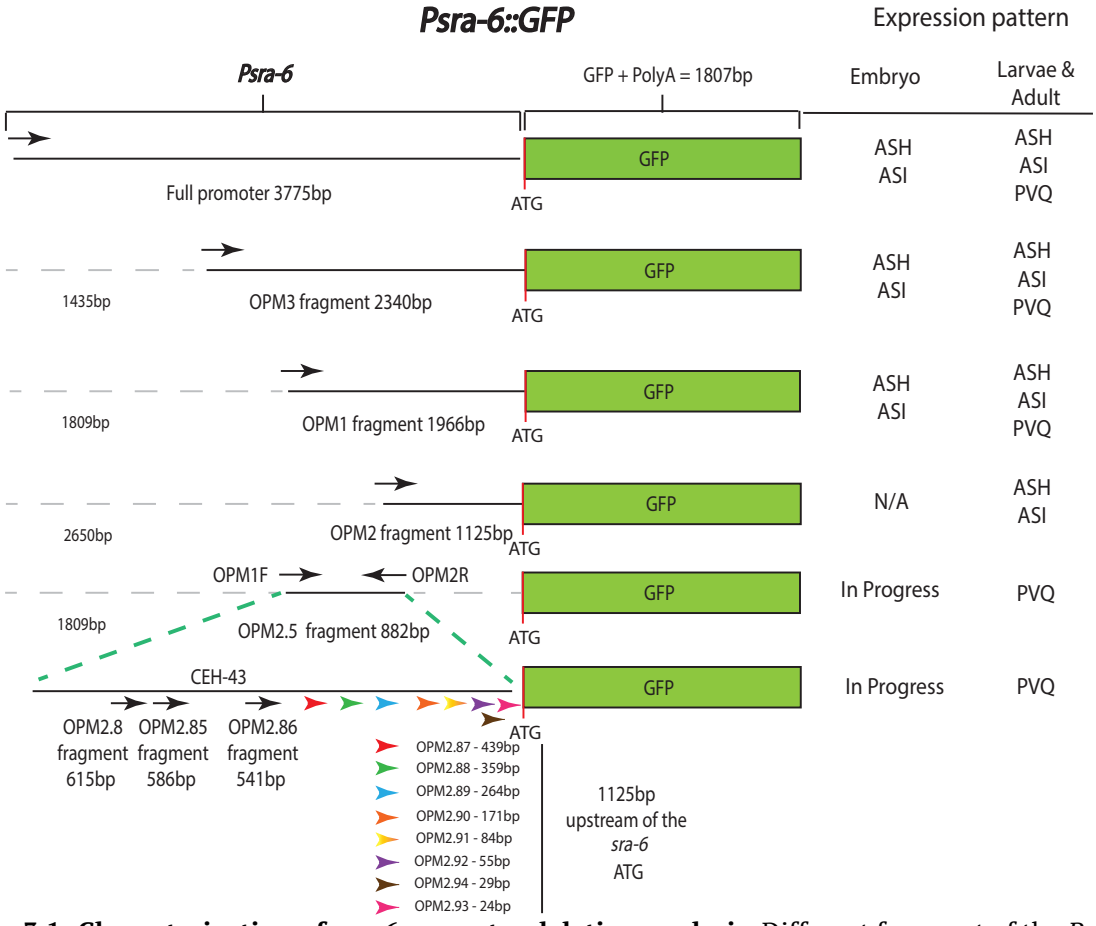

Figure 3.1. Characterization of *sra-6* promoter deletion analysis. Different fragment of the *Psra-6* were obtained by PCR (OPM3 to OPM2.86) and subsequently used to generate transgenic reporter lines in wild-type worms. Small colour arrows represent shorter fragments of the OPM2.5 fragment. N/A – not available; Number of lines analysed – 1 to 2; n>20, animals analysed at the L4 stage.

Analysis of transgenic line revealed that the OPM3 fragment is able to drive expression in an extra set of neurons near the nerve ring (Figure 3.2c to h). This finding reveals a possible regulatory region upstream of the OPM3 primers that controls the expression of this reporter in this set of neurons (Figure 3.2c to h).

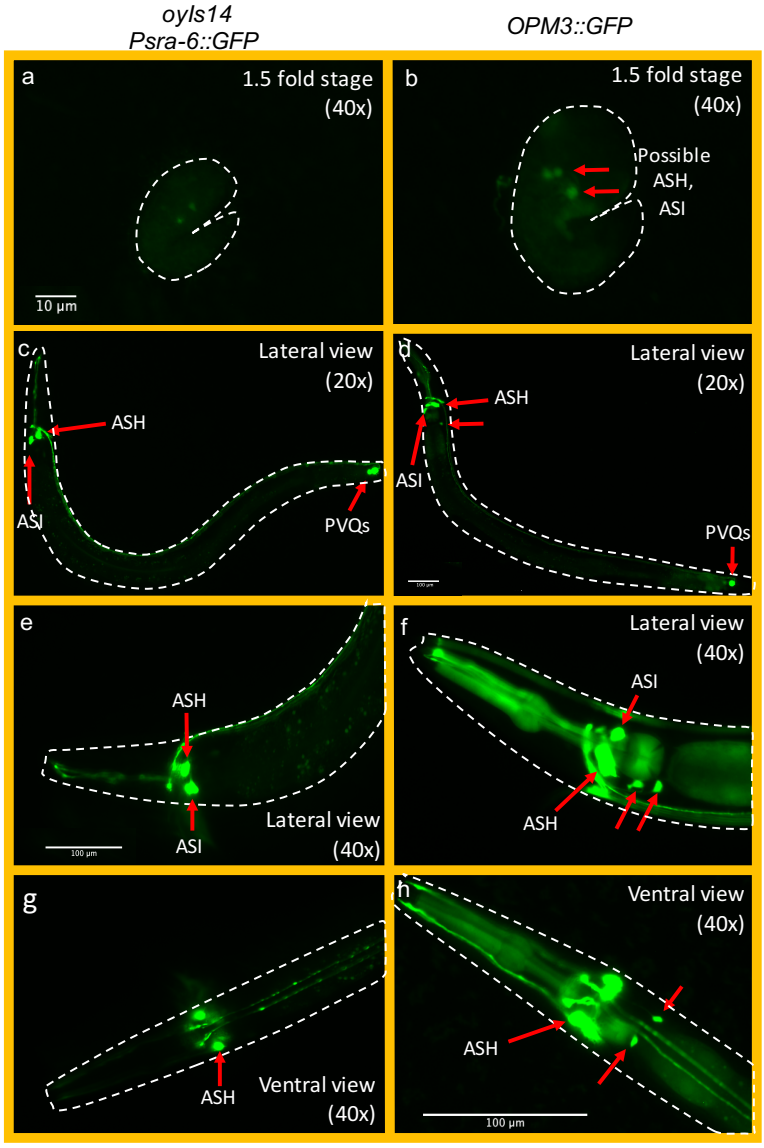

Figure 3.2. Characterization of *Psra-6* and expression pattern. Transgenic worms expressing GFP under the control OPM3 fragments compared with the control *Psra-6::GFP* transgenic worms. **A –** Single egg expressing *Psra-6::GFP*; **B –** Single egg expressing *OPM3::GFP*; **C** – Lateral view of a transgenic worm expressing *Psra-6::GFP*; **D –** Lateral view of transgenic worm expressing *OPM3::GFP*. Red arrows point an extra set of neurons; **E –** Lateral view of the head of a worm expressing *Psra-6::GFP*; **F –** Lateral view of the head of a worm expressing *OPM3::GFP*. Red arrows point an extra set of neurons; **G –** Ventral view of the head of a worm expressing *Psra-6::GFP;* **H –** Lateral view of the head of a worm expressing *OPM3::GFP*. Red arrows point an extra set of neurons. Scale bar – A, B (10 µm); C, D (100 µm); E, F, G, H (100 µm). Worms imaged at L4 and young adult stage (c to h). Number of lines analysed – 1 to 2.

Moreover, analysis of transgenic lines revealed that the OPM3 and OPM1 fragments drove stronger GFP expression at the 1.5-fold embryonic stage (Figure 3.2a, b & 3.3d) when compared with the control, however, these strains still presented GFP expression in the ASI, ASH and PVQ after embryogenesis. In contrast, worms injected with the OPM2 fragment did not present GFP expression in the PVQ neurons but maintained GFP expression in the ASI, ASH and in additional structures. These data suggest that the region responsible for controlling the expression of *sra-6* in the PVQs is located between the OPM1 and OPM2 primers. To test this hypothesis, I amplified a fragment between OPM1 and OPM2 primers to drive GFP (named OPM2.5) and cloned this fragment into to a GFP expression plasmid. The analysis of the OPM2.5 transgenic worms revealed that the GFP expression was only present in the PVQs (Figure 3.3b). This result confirms that the regulatory element responsible for *sra-6* expression in the PVQs is located in this region. However, almost all of these transgenic animals exhibit defects in PVQ development, such as outgrowth defects and midline crossovers (PVQ defect example in Figure 3.3c).

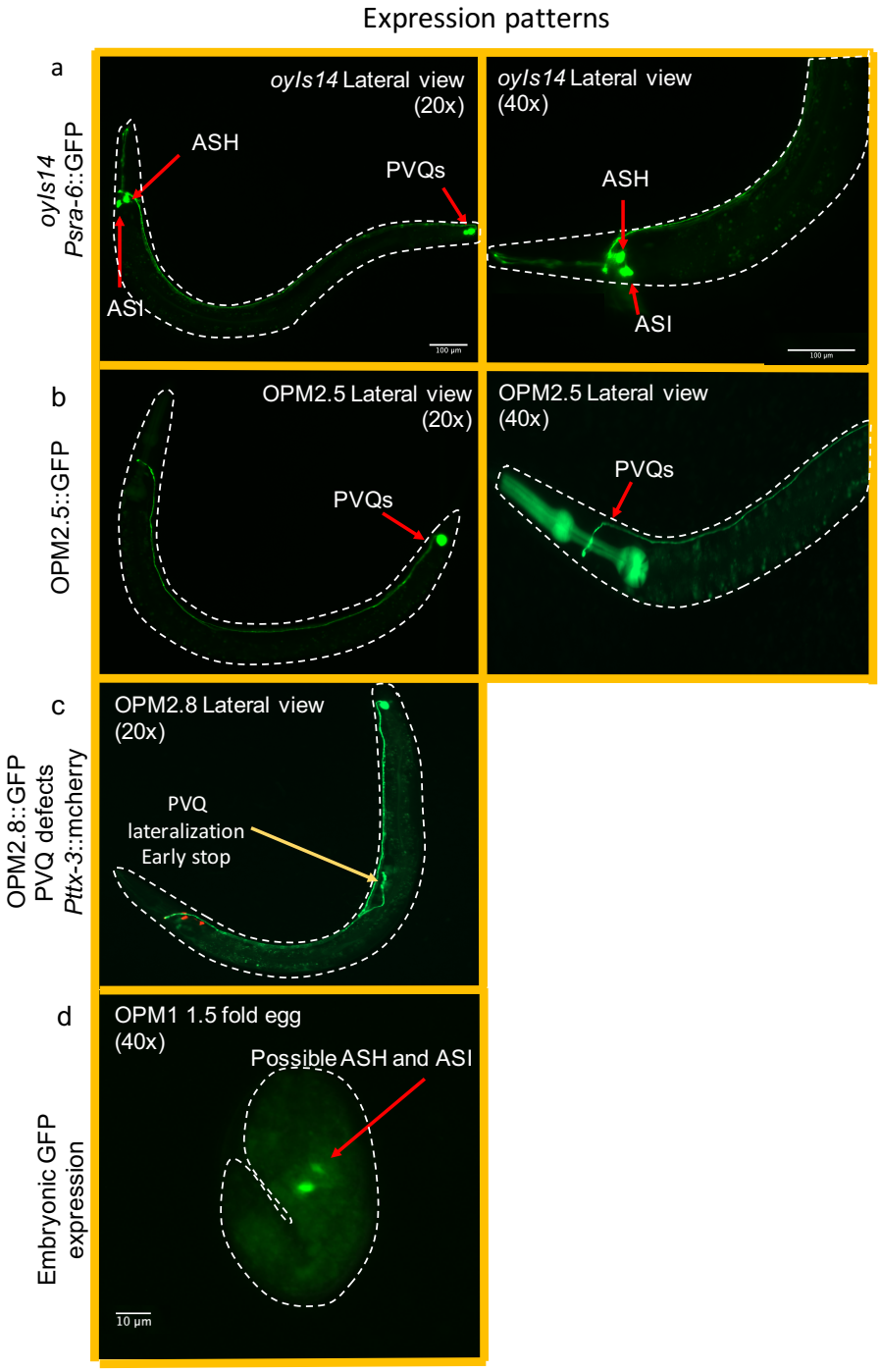

Figure 3.3. Characterization of *Psra-6* and expression pattern. Transgenic worms expressing GFP under the control of different *Psra-6* fragments. **A –** Control strain (*oyIs14*) expressing GFP under the control of the full *sra-6* promoter. **B –** Transgenic worms expressing GFP under the control of OPM2.5 fragment. In this strain GFP is only expressed in the PVQ neurons. **C –** Transgenic worms expressing GFP under the control of OPM2.8 fragment. These worms still present GFP expression only on the PVQ and present PVQ guidance defects (yellow arrow). **D –** Single egg from OPM1 strain expressing GFP, possibly in the ASI and ASH neurons due to the cell location (red arrow). Scale bar – A to C (100 µm); D (10 µm).

This interesting finding suggests that a quenching phenomenon may be occurring, i.e., an important factor for PVQ development is being depleted by overexpression of the *sra-6* promoter. To better understand this problem, I quantified the PVQ defects for the OPM fragments used (Figure 3.4).

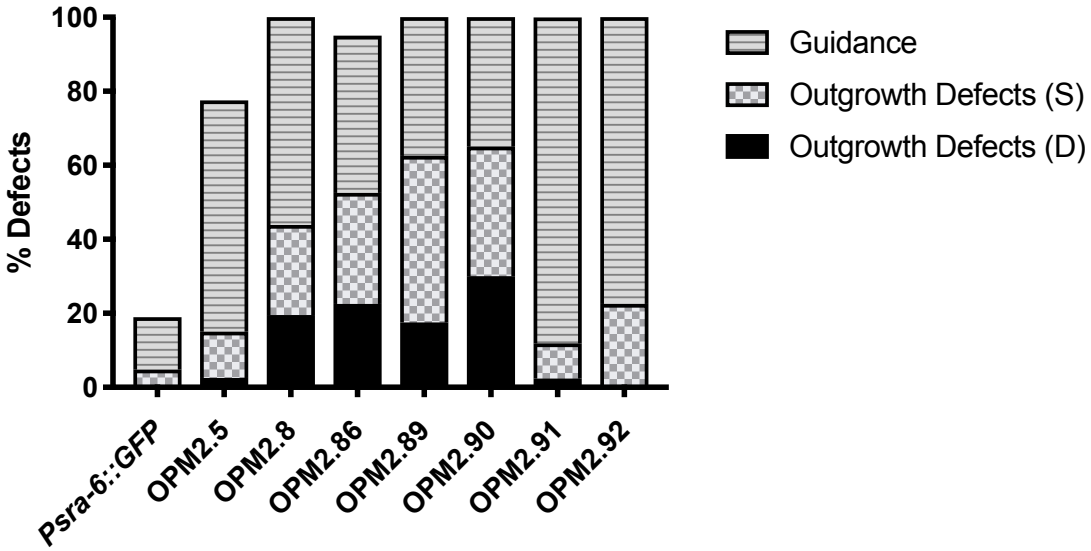

Figure 3.4. Quantification of the PVQ developmental defects in transgenic worms injected with the different OPM fragments when compared with the *Psra-6::GFP*. Percentage PVQ guidance and outgrowth defect present in the different OPM transgenic worms. PVQ defect were categorized into guidance and outgrowth defects (S stands for single PVQ outgrowth defect and D stand for double PVQ outgrowth defect), n>22 worms.

The graph in the figure above demonstrates that less that 20% of the transgenic worms injected with the full *sra-6* promoter present mild PVQ guidance defects. In contrast, worms injected with smaller fragments present more that 70% PVQ defects. For example, worms injected with OPM2.5 fragment present a higher percentage of guidance and outgrowth defects (60% and 15% respectively) when compared to the full promoter transgenic strain. Moreover, worms injected with even smaller fragments, such OPM2.8, OPM2.86, OPM2.89 and OPM2.90, present 20 – 50% in PVQ outgrowth defects when compared to the OPM2.5 strain. Finally, by further reducing the OPM fragment size, OPM2.91 and OPM2.92, outgrowth defect reduced to OPM2.5 levels. A possible explanation for these results is the presence of important regulatory regions in the OPM2.5 fragment that is required for PVQ development (Figure 3.5).

The overexpression of this small fragments of the *sra-6* promoter (OPM2.5 to OPM2.92) may lead to the transcriptional dysregulation of genes required for the correct PVQ development (Figure 3.5a).

Since the PVQ defects were always present in all the strains generated in this work. I performed a bioinformatics analysis on the OPM2.5 sequence to identify possible transcription factor binding sites. This analysis revealed a few predicted TF binding sites in this fragment (Figure 3.5b), such as for the CEH-43 homeodomain transcription factor (Figure 3.5b) that is expressed in PVQs and known to be involved in neuronal development of dopaminergic neurons (Aspock and Burglin, 2001; Doitsidou *et al.*, 2013).

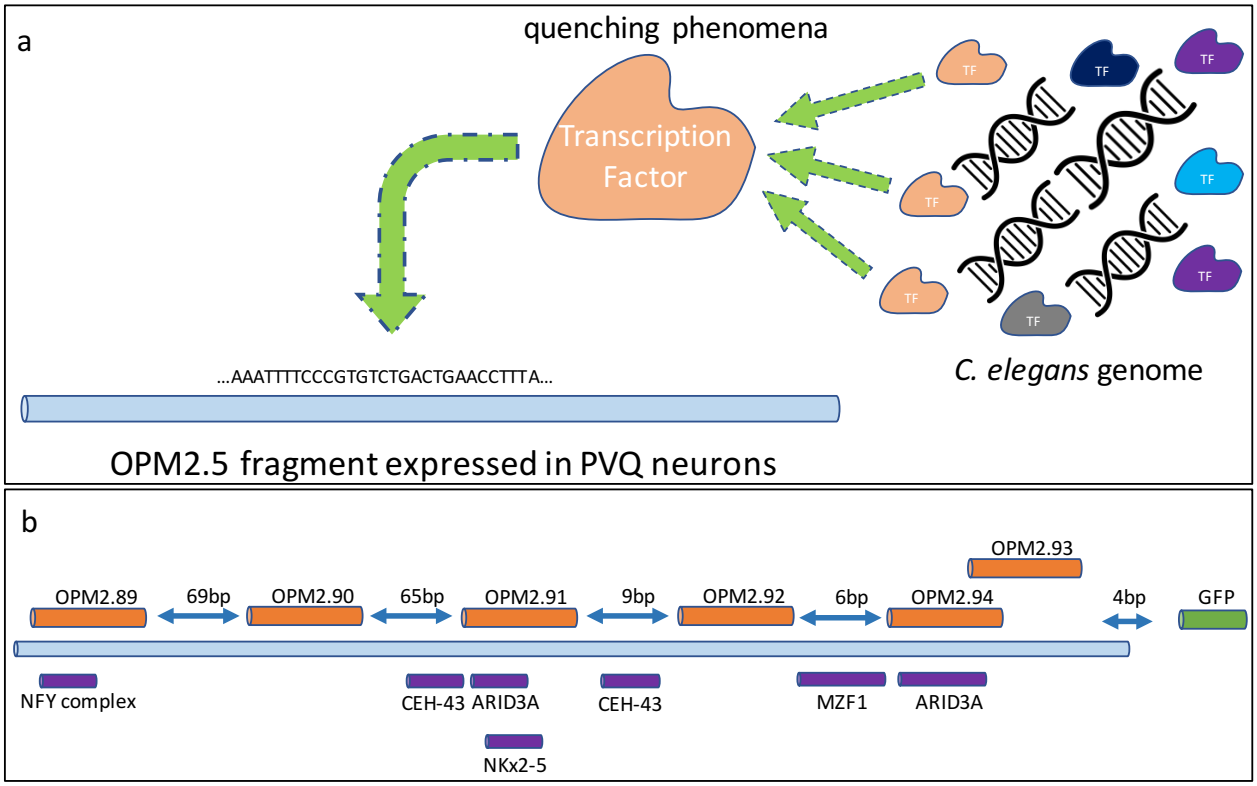

Figure 3.5. Schematic representation of the quenching phenomena affecting the promoter characterization experiments and the location of the putative binding sites for the different TF. **A –** The images represent a TF being quenched by being directed to the OPM2.5 fragment leading to the PVQ defects. **B –** This scheme represents the binding sites present in a smaller fragment (OPM2.89F) that presents 100% of PVQ defects. The most prominent candidates that could cause these defects are the NFY complex, CEH-43, Arid3a (CFI-1 in *C. elegans*), NK2-5 (CEH-22 in *C. elegans*) and MZF1. The orange cylindrical shapes - OPM2.89 to OPM2.93, represent the binding site for the primers used.

To date, no link between *ceh-43* and PVQ development or maintenance has been revealed. To determine whether CEH-43 may be important for PVQ development, I injected two smaller fragments with and without the CEH-43 predicted binding site (OPM2.8 and OPM2.86, respectively). The results revealed that both OPM2.8 and OPM2.86 lines presented GFP expression in the PVQs. However, these worms still presented PVQ developmental defects. Also, further analysis of *ceh-43(ot406)* hypomorphic mutant worms crossed with a PVQ reporter did not present guidance or outgrowth defects, and therefore does not support our hypothesis. These findings suggest that the region responsible for controlling *sra-6* expression in the PVQs, with PVQ specification elements, is located downstream of the OPM2.86 primer.

Another candidate was the CFI-1 (ARID3A in mammals) transcription factor previously reported to be required for neuronal development in *C. elegans* (Shaham and Bargmann, 2002). To assess if the CFI-1 TF was responsible for the defects observed, I reduced the size of the OPM2.5 fragment in order to remove the putative CFI-1 binding sites (Figure 3.5b) (OPM2.89 has a binding site upstream and OPM2.92 fragment also removes another binding site), but again with no positive results (Figure 3.4). Further analysis of *cfi-1(ky651)* mutant worms crossed with a PVQ reporter did not present any detectable guidance or outgrowth defects, and therefore does not support our hypothesis.

The analysis of the OPM2.5 fragment also revealed more predicted binding sites downstream of the OPM2.86 primer, such us the binding site for NFY, JUN-1, FOS-1, MZF1 (*Y48G8AL.10* and *Y55F3AM.14* in *C. elegans*). I conducted a screen for axon guidance defects, analysing PVQ development for some these candidate factors. The preliminary results obtained indicated that one of these candidates – the NFY transcription factor – is involved in PVQ development. To support this finding, a forward genetic screen, independently carried out using a *Psra-6::GFP* reporter strain by the Pocock laboratory, identified the NFY transcription factor as a possible candidate to be required for PVQ development.

Since it is not clear what causes the PVQ developmental defects in animals expressing the OPM2.5 fragment, I cloned the OPM2.5 fragment into two different vectors with different 3' UTRs. Amazingly, transgenic lines for both of these vectors did not present any detectable GFP expression in any cell. Based on preliminary analysis, I hypothesized that the cause of the guidance and outgrowth defect are due to a concerted action of the OPM2.5 fragment and the *let-858* 3'UTR from the plasmid used during the majority of this work. To further reinforce this hypothesis, I injected worms expressing GFP in the PVQs with only the OPM2.5 fragment and analysed PVQ development. Interestingly, no PVQ developmental defects were observed in these transgenic lines. Moreover, I injected the original plasmid, containing the *let-858* 3'UTR, and lacking the OPM2.5 fragment, and did not detect PVQ defects. These preliminaries observations data support our hypothesis that both OPM2.5 and *let-858* 3' UTR work together to disrupt PVQ development possibly by disrupting the transcriptional machinery required for PVQ development. Further analysis is still required in order to understand if the *let-858* 3' UTR alone is able to cause PVQ development defects.

3.5. Discussion

3.5.1. Promoter *sra-6* regulatory motifs

In this work, I performed a genetic and cis-regulatory analysis to identify motifs within the *sra-6* promoter that are responsible for expression in different neurons. My specific aim was to identify the regulatory region responsible for PVQ gene expression. Here, I found evidence that the 3' end of the *sra-6* promoter (OPM2 fragment) is responsible for controlling gene expression in two pairs of neurons in the head of the worm (ASH and ASI neurons). The overexpression of small fragments of the *sra-6* promoter (OPM2.5 to OPM2.92) is enough to drive GFP expression specifically in the PVQs but may lead to the transcriptional dysregulation of genes required for the correct PVQ development (Figure 3.5a). To overcome this limitation, I reduced the concentration of the injected constructs used in this chapter but was not able to remove the guidance defects. Nevertheless, by dissecting the *sra*-6 promoter I identified the promoter region (OPM2.5 fragment) responsible for controlling *sra-6* expression specifically in the PVQs. Furthermore, this promoter region contains predicted transcription factor binding motifs that could be relevant for the control of *sra-6* expression in the PVQs. Thus, I analysed genetic mutants for specific transcription factors predicted to bind to the OPM2.5 region such as CEH-43, CFI-1 and the NFY complex transcription factors.

The CEH-43 TF is a good candidate to be involved in PVQ development since it is expressed in the PVQs during development and regulates the neuronal fate of other neurons in *C. elegans* (Aspock and Burglin, 2001; Doitsidou *et al.*, 2013). Previous studies conducted by Doitsidou and colleagues demonstrated that *ceh-43* gene, an orthologue of the fly Distalless and vertebrate Dlx homeobox genes, is required for the correct dopaminergic neuronal fate acquisition in *C. elegans* (Aspock and Burglin, 2001; Doitsidou *et al.*, 2013). Mutation of this gene leads to downregulation of dopaminergic specific markers, such as the *dat-1* gene. However, analysis of mutant worms with a missense mutation on the *ceh-43(ot406)* gene did not reveal any detectable PVQ developmental defects such guidance and outgrowth defects. This

leads me to conclude that the phenotype observed previously in my analysis is likely due to other factors.

Another possible candidate was the CFI-1 transcription factor, an orthologue of ARID3A, ARID3B and ARID3C in vertebrates. CFI-1 is required for URA and IL2 differentiation in hermaphrodite worms by preventing the URA and IL2 neurons from expressing CEM-specific male genes (Shaham and Bargmann, 2002). Moreover, CFI-1 promotes the correct formation of the glutamate receptor (GLR-4) in the URA, AVD and PVC neurons, which is important for the correct function of these neurons (Shaham and Bargmann, 2002). However, this approach was unsuccessful and the analysis of *cfi-1(ky651)* mutant worms crossed with PVQ::GFP reporter did not present any detectable guidance or outgrowth defects.

Finally, we focused our attention in the third candidate for PVQ regulation - the NFY TF complex. This complex is important for post-mitotic neuron maintenance in mice and it is involved in neuronal type-dependent pathologies (Yamanaka *et al.*, 2008; Yamanaka *et al.*, 2014; Yamanaka *et al.*, 2016). Moreover, results from a forward genetic screen conducted in our lab found that the *nfyc-1* gene, encoding one of the NFY complex subunits, is required for PVQ development. This result suggests that the NFY complex may be required for correct PVQ development, and as such, account for the defects observed so far in this work.

Preliminary results showed that the NFY transcription factor complex is essential to control *Psra-6*::GFP reporter expression in the PVQs. Despite not being analysed in detail, based on previous studies I predicted that either the NFY complex may directly control *sra-6* expression in the PVQs by binding to the CCAAT binding site present in the promoter, or indirectly through interactions with other co-activators of transcription (Deng *et al.*, 2007). Another possibility, is that the NFY complex affects the expression of *sra-6* indirectly by regulating other transcription factors required for PVQ development. Further analysis regarding the role of the NFY complex and possible co-activators will be described in the next chapters.

In summary, since very little is known about the gene regulation required for PVQ development, function and maintenance. This work presents the perfect opportunity to discover more about PVQ biology.

3.6. Future directions:

I aim to screen other transcription factor mutants predicted to bind the *sra-6* fragment and to continue the studies on *sra-6* fragment UTR in order to find the guidance and outgrowth defects origin. I will analyse mutant worms for the following transcription factors: *Y48G8AL.10* and *Y55F3AM.14 (*MZF1). I will also proceed with the generation of transgenic worms for the *let-858 3'UTR* in order to test its role in the guidance and outgrowth defects observed.

Chapter 4. Dissecting the role of the NFY transcription factor complex in nervous system development

4.1. Abstract

The establishment of correct brain architecture during development is an exceptionally complex process requiring precisely controlled neuron specification, neuronal migration and axo-dendritic outgrowth and guidance. These events are controlled by multiple conserved transcription factors, guidance systems and environmental cues. The *Caenorhabditis elegans* nervous system is an excellent model to study brain development due to its relative simplicity and conserved nature of development.

My PhD project aims to identify molecular mechanisms that drive brain development. Using an unbiased genetic screen, we identified a Nuclear Factor Y transcriptional complex (NFY) that controls the development of a pair of glutamatergic interneurons.

The NFY family is one of the most abundant and conserved transcription factors in eukaryotes and is involved in the regulation genes associated with developmental steps, such as tail patterning C. *elegans* and nervous system maintenance in mammals. NFY is a trimeric complex composed of NFY-A, -B and -C subunits that regulate gene expression through binding specific motifs in promoter regions.

My data shows that the NFY complex regulates neuronal fate and axon guidance of specific neurons. Using single-cell resolution analysis, transcriptomics and ChIP sequencing, I will decipher the molecular mechanism(s) through which NFYs control neuronal development.

4.2. Introduction

Correct establishment of nervous system architecture during development is a complex process that requires precise neuronal specification, neuronal migration and axon/dendrite extension to connect with specific target cells. Transcription factors act by controlling a specific repertoire of genes essential for neuronal development. Gene expression is precisely controlled by DNA sequences located in the promoter region located upstream of the transcriptional start site of target genes. These DNA regions possess several conserved motifs, such as TATA, GC and CCAAT boxes, which are recognized by transcription regulators that enhance or repress transcription (Tjian and Maniatis, 1994). The NFY transcription factor is one of the most common and abundant transcription factors in eukaryotes. It is involved in the regulation of genes, like the *egl-5*/HOXB, and its dysfunction is implicated in different diseases, such as cancer, neuronal degeneration and Huntington's disease (Benatti *et al.*, 2016; Deng *et al.*, 2007; Yamanaka *et al.*, 2008; Yamanaka *et al.*, 2014; Yamanaka *et al.*, 2016). The NFY transcription factor forms a heterotrimeric complex composed of three different subunits, NFYA-I, NFYB-I and NFYC-I. These subunits interact with each other in a specific manner through evolutionary conserved motifs, called HAP2, HAP3 and HAP5, respectively. HAP3 and HAP5 possess conserved histone-folding motifs (HFM), which are composed of three or four ∂-helices (∂1, ∂2, ∂3 and ∂N/C) (Baxevanis *et al.*, 1995). Functional studies showed that these conserved HFMs interact with each other in a head to tail orientation – ∂1 from NFYC interacts with ∂3 from the NFYB – to form dimers. The NFYC/NFYB heterodimer forms a complex surface that recruits the NFYA subunit to form a heterotrimeric complex (Mantovani, 1999). Previous studies indicate that the NFYA subunit possesses a DNA binding domain recognizes a specific motif in the promoter of target genes – a consensus CCAAT box, which is one of the most frequent and conserved regulatory motifs present in eukaryotic promoters.

Recently, it was found that the interaction between NFY and the promoter of the HSP-70 gene, which is known to help reduce the toxicity of polyglutamine stretches in Huntington's disease, can be disrupted due to an interaction between mutant

Huntingtin protein (Htt) and the polyQ region within the NFYA and NFYC proteins (Yamanaka *et al.*, 2008). This study revealed that NFYA and NFYC subunits are sequestered by Htt, both *in vitro* and in an Htt mutant mouse model, promoting transcriptional dysregulation and triggering neurodegenerative disease (Yamanaka *et al.*, 2008). Additional studies carried out by Yamanaka revealed that the NFY complex is also important in specific postmitotic neuronal to prevent degeneration in the mouse brain (Yamanaka *et al.*, 2014). However, despite its ubiquitous expression, the function of the NFY complex in the nervous system is poorly understood.

An unbiased forward genetic screen performed by the Pocock Lab identified a point mutation in the *nfyc-1* gene that causes developmental defects in the *C. elegans* PVQ interneurons. In this study, I aimed to determine the precise role of the NFY complex in neuronal development. I analysed the expression of terminal PVQ fate reporters that represent readouts of PVQ identity. In this analysis, I found that mutations in the each of the NFY complex subunits cause defects in PVQ specification and axon guidance. In addition, I found that the NFY complex functions cell-autonomous to control PVQ specification. Moreover, I found that mutations in CCAAT sites, that may be bound by NFYA-1, cause reduced expression of PVQ terminal fate reporters. Interestingly, both NFYA-1/2 subunits control PVQ fate but only NFYA-1 is involved in PVQ guidance. Finally, I found that the NFY complex also controls expression of general neuronal features. Finally, I aim to perform mass spectrometry in other to screen for transcriptional co-activators that work together with the NFY complex during PVQ development.

4.3. Methods

Strains

All strains were maintained at 20°C for two generations prior to scoring as previously described (Brenner, 1974), unless stated otherwise. All the transgenic and mutant strains used in this work are detailed in Table 2.1 and Supplemental Table 2.

Genetics

The following mutant strains were used in this work: *nfya-1(ok1174), nfya-1(bp4), nfya-2(tm4194), nfyb-1(cu13), nfyb-1(tm4257), nfyc-1(tm4145), nfyc-1(rp120)* (Supplemental Table 2). The mutant strains used in this work represent strong loss-of-function alleles as the majority of each protein is deleted (Figure 4.1a). All the mutant and transgenic strains used in this chapter are depicted in Supplemental Table 2.

Expression constructs and Transgenic animals

These reporter strains were already published and were used in this study (Supplemental Table 2).

Fluorescence microscopy

GFP expression and neuron morphology was scored in L1 larvae and L4 hermaphrodites anesthetized in 10 mM sodium azide (NaN_3) and mounted on 5% agarose on glass slide. Images were capture using an automated fluorescence microscope (Zeiss, AXIO) and ZEN software. Fluorescence levels were measured using FIJI software (ImageJ V.1.47n).

Scoring of crossover defects

The guidance or outgrowth defects of PVQs were classified into categories: guidance defect occurring due to misguidance of the PVQL or PVQR, single or double stop for outgrowth defect (PVQL or PVQR).

Statistical analysis

Statistical analysis was performed in GraphPad Prism 7 using t-student test, One-way ANOVA for multiple comparison. Values are expressed as mean – standard error of the mean (SEM) or mean – standard deviation (SD). Differences with a *p value<0.05 were considered significant.

4.4. Results

4.4.1 Characterising the NFY transcription factor role in PVQ development

In the previous chapter, results from the bioinformatics analysis and results from a previous forward genetic screening conducted in our laboratory suggested that the NFY complex could play a role in PVQ development. To test this hypothesis, and to gain insight on the NFY family's function during PVQ development, I analysed loss of function alleles and performed co-localization studies and analysis of PVQ terminal fate reporters to dissect the precise function of the NFY complex in the PVQ neurons - detailed below (Figure 4.1a and Figure 4.1b).

To confirm that the NFY transcription factor subunits are expressed in the PVQ neurons, I performed co-localization experiments with a PVQ specific *gfp* reporter and three transgenic strains (*Pnfya-1*::*H1-mCherry, Pnfyb-1*::*H1-mCherry* and *Pnfyc-1*::*H1-mCherry*). Our analysis revealed that the NFY complex subunits, -A, -B and –C are ubiquitously expressed in *C. elegans* throughout development. The co-localization studies revealed that all NFY subunits are expressed in the PVQ neurons (Figure 4.1b). This reinforces our hypothesis that the NFY transcription factor is required to regulate gene expression in the PVQs. Further experiments will be performed to show that *nfy* gene expression is required cell-autonomously to control PVQs development.

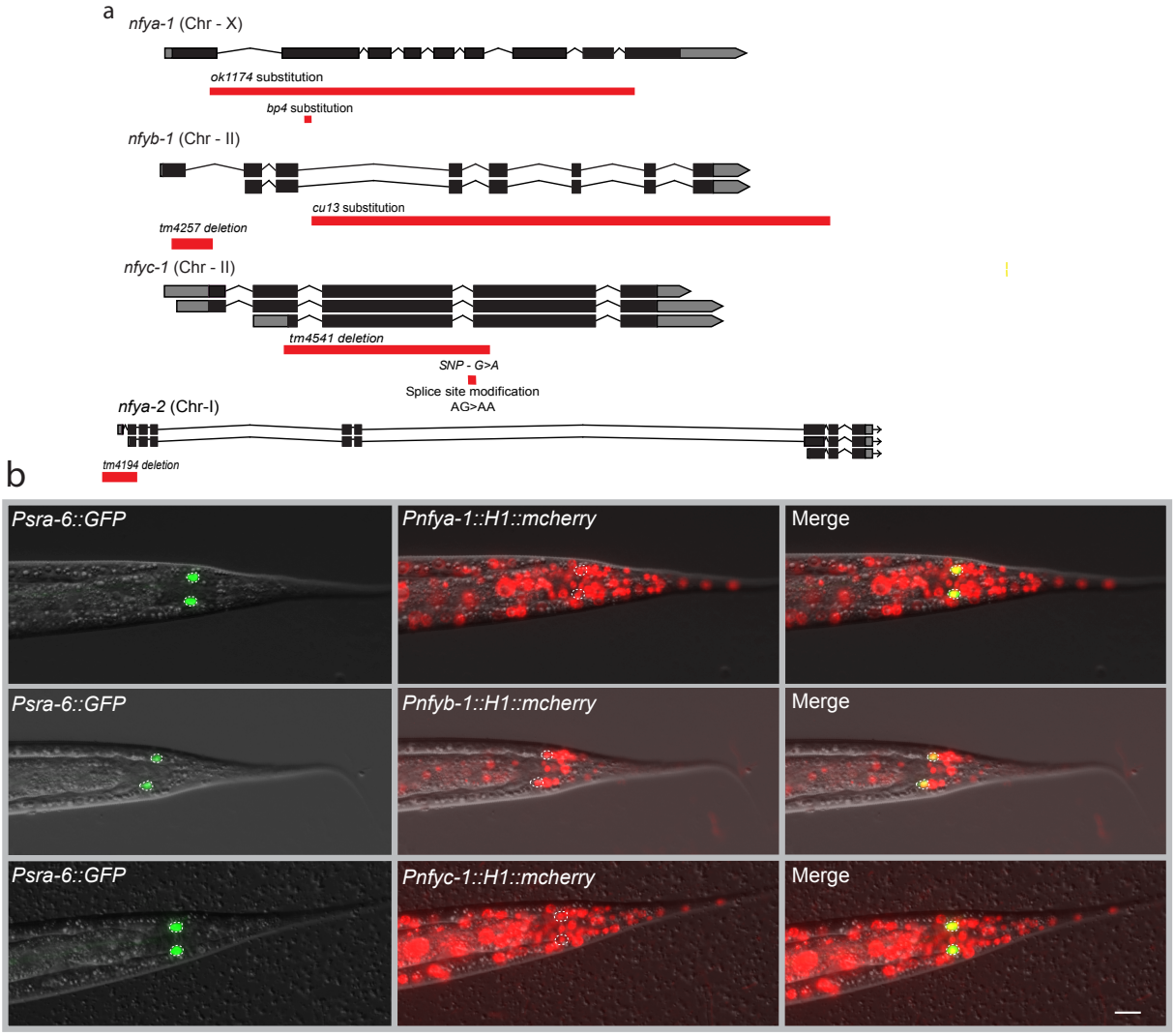

Figure 4.1. Co-localization studies for NFYA-1, NFYB-1 and NFYC-1 subunits in PVQ neurons at L4 stage. A - Molecular identity of mutant alleles for *nfya-1(ok1174), nfya-1(bp4), nfya-2(tm4194), nfyb-1(cu13), nfyb-1(tm4257)* and *nfyc-1(tm4541), nfyc-1(rp120)*. The *nfya-1*(*ok1174)* allele is a substitution allele, substituting 1837bp for two thymine residues (between exon1 and exon 9). The *nfya-1(bp4)* allele is a single nucleotide polymorphism (SNP) mutation that leads to the formation of a stop codon in residue 102aa. The *nfya*-2 *(tm4194)* is a deletion allele, deleting 360bp in 3' region of the promoter and exon 1. The *nfyb-1*(*cu13)* allele is a substitution allele, substituting 5117bp for one thymine residue (downstream to exon 3). The *nfyb-1(tm4257)* allele is a deletion allele, deleting 355bp in exon 1. The *nfyc-1(tm4541)* allele is a deletion allele, substituting 417bp (between exon 2 and exon 4). The *nfyc-1(rp120)* allele is a SNP mutation at the splicing site of exon 4 that likely leads to a truncated protein. All strains represent strong loss-of-function alleles as the majority of each protein is deleted. **B** - Co-localization for the NFYA-1, NFYB-1 and NFYC-1 subunits (red) was performed at the L4 stage with the help of a transcriptional transgenic markers (see Supplemental Table 2). These experiments revealed that both NFYA-1, NFYB-1 and NFYC-1 subunits are expressed in the nucleus of the PVQ neurons at L4 stage due to its co-localization with the PVQ reporter (green). These results reinforce our hypothesis that NFY complex is required for the correct PVQ development. Scale bar (10 µm).

In order to characterize the role of the NFY TF complex in PVQ development, I used mutant strains for each NFY subunit. I noticed that *nfy* mutant worms exhibit defects in PVQ::GFP expression, where no GFP expression is detectable in one or both PVQs, referred in this work as "PVQ off" (Figure 4.2a and Figure 4.2b).

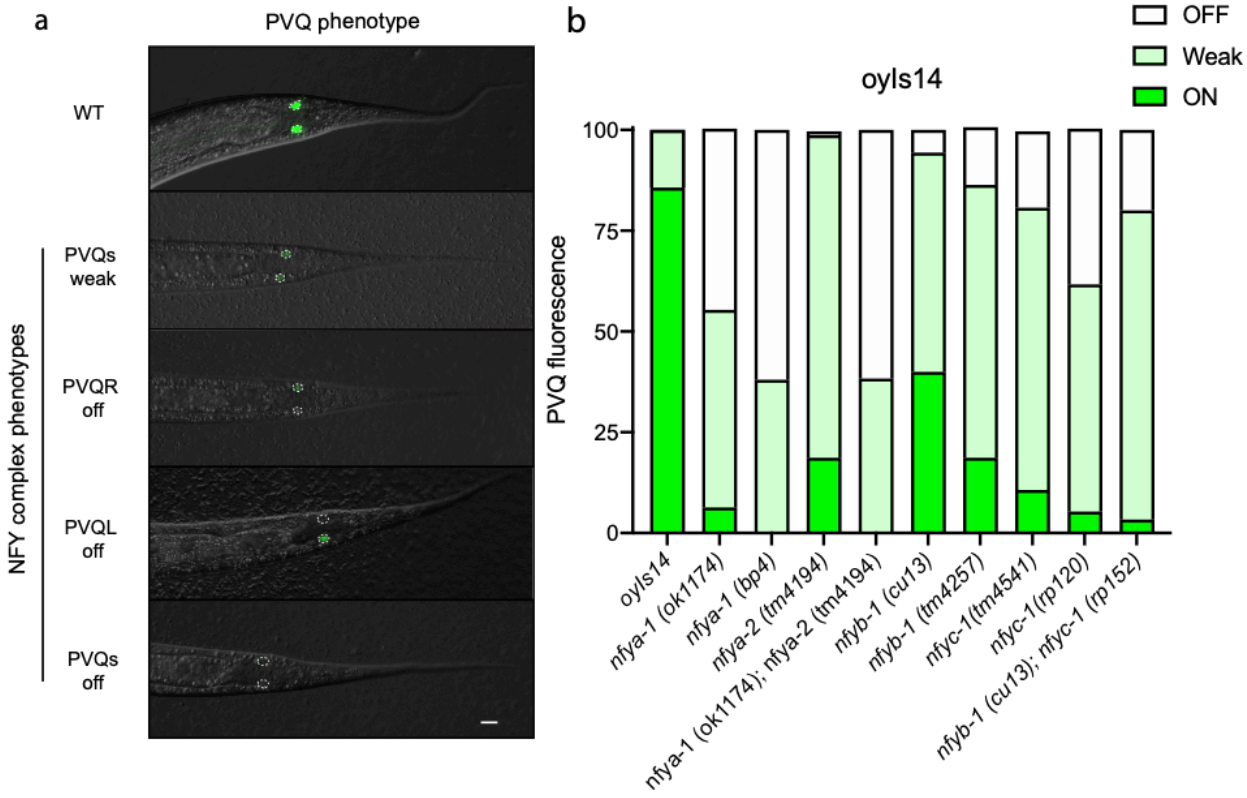

Figure 4.2. Characterization of *nfya-1, nfyb-1 and nfyc-1* genes during PVQ development. Two independent mutant alleles for each *nfy* gene (with exception of *nfya-2*) were crossed with a PVQ::GFP reporter strain (*oyIs14*) that expresses GFP in the PVQ neurons, and subsequently scored for PVQ developmental defects. The expression of *oyIs14* reporter is not affected (100% ON) in the ASH and ASI neurons in the different mutant strains used. **A –** Representative images of phenotypes associated with the NFY complex – PVQs weak, PVQR/PVQL or both PVQs off. **B –** Percentage of neurons that do not show GFP expression for each mutant NFY subunit. Data is expressed as mean ± SEM and statistical significance was assessed using one-way ANOVA with multiple comparison with Dunnet's multiple comparison test. ns, not significant, n=3 replicates (60 - 160 neurons). Scale bar (10 µm).

Here, I found that *nfyb-1* mutant worms present a smaller percentage of PVQ off (around 10% to 20%) when compared to the *nfyc-1* with 20%-40% and the *nfya-1* with 70% PVQ off (Figure 4.2b). Based on this result, I hypothesized NFYA-1 may play a

partial independent role from the NFYB-1 and NFYC-1. To test this hypothesis, I used CRISPR-Cas9 to introduce a premature nonsense codon into the first exon of *nfyc-1* in the *nfyb-1(cu13)* mutant background since both gene are on chromosome II. The *nfyb-1; nfyc-1* double mutant exhibits similar penetrance of *sra-6p::GFP* expression in the PVQs as *nfyb/c-1* single mutants (Figure 4.2b). These data suggest that NFYA-1 and/or NFYA-2 may partially regulate *Psra-6::GFP* expression independently of the NFYB/C subunits. Additionally, it is interesting to note that *C. elegans* is the only model organism with a NFYA-2 subunit, and that mutant worms for the NFYA-2 subunit do not exhibit any PVQs off but present PVQs with weak GFP expression (Figure 4.2b). This data suggests that the NFYA-1 and NFYA-2 may have different roles. Moreover, NFYA-1 and NFYA-2 double mutant worms present a 16% increase in PVQs off when compared with the *nfya-1 (ok1174)* single mutant worms. This suggests that NFYA-2 acts in a separate NFY complex to control PVQ fate, with the NFYA-1 subunit playing a stronger role. Another possible explanation is that the mutation analysed for the *nfya-2* gene may not affect the smaller isoform of the NFYA-2 protein, allowing the *nfya-2* gene to maintain a certain level of functionality.

To confirm that the loss of PVQ terminal fate reporter expression is due to the absence of NFY function and not due to background mutations, I performed transgenic rescue analysis. For this experiment, I focused on *nfya-1* mutant worms, since they presented the most severe phenotype, and injected them with a *nfya-1* fosmid, i.e, a wild type region of the DNA (containing the genomic region of the *nfya-1* gene) cloned into a plasmid. This approach aims to re-supply a wild type copy of the *nfya-1* gene to mutant worms, thus rescuing the phenotype. The result presented in Figure 4.3a show complete rescue of the PVQ off phenotype. Next, I asked if it was possible to cell autonomously rescue this phenotype. To this end, I designed a construct where I used a PVQ specific *npr-11* promoter to drive the expression of *nfya*-1 cDNA and found that this also rescued the *nfya-1(ok1174)* phenotype (Figure 4.3b).

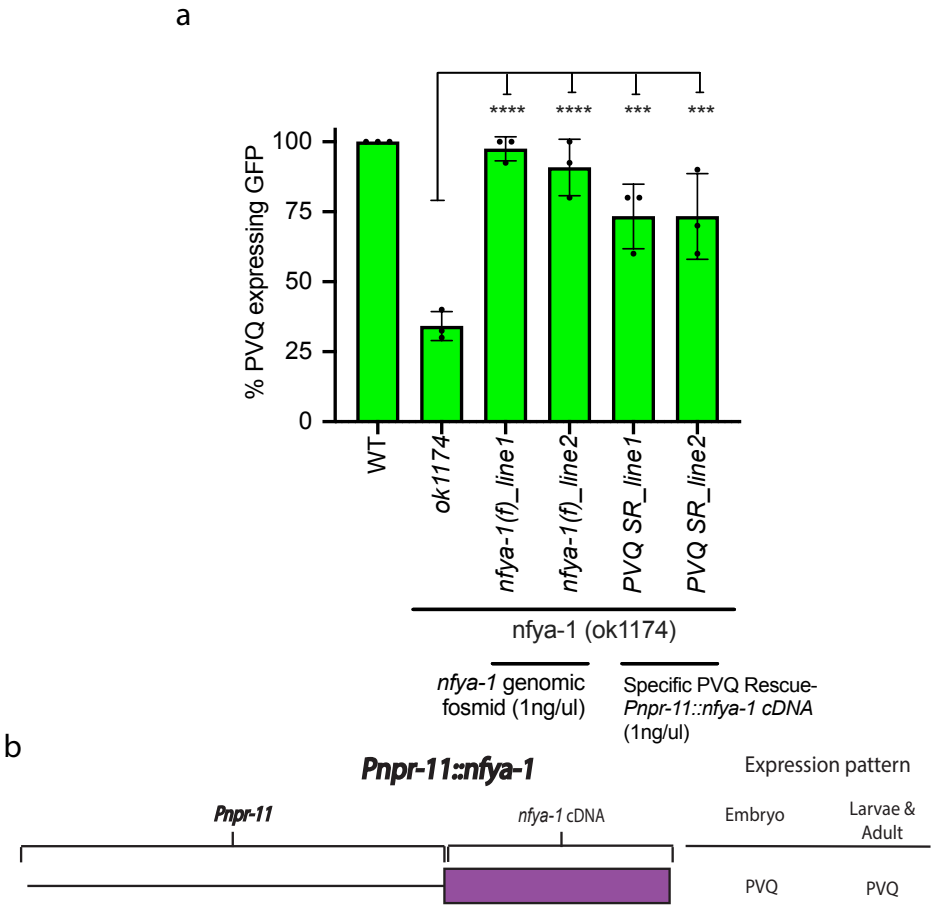

Figure 4.3. Rescue experiments for the *nfya-1* mutation. Mutant worms for the *nfya-1* gene were crossed with an PVQ GFP reporter (*oyIs14*) and were injected with a fosmid (genomic) or PVQ specific rescue construct in order to obtain two independent lines for each condition. **A** – Fosmid and the PVQ-specific expression of *nfya-1* rescues the *nfya-1* mutant phenotype. **B** – Schematic representation of the PVQ specific rescue construct [*Pnpr-11::nfya-1 cDNA]*. Data is expressed as mean ± SEM and statistical significance was assessed using one-way ANOVA multiple comparison Dunnett's test. ***p<0,001, ****p<0,0001. n= 3 replicates (15 worms per replicate).

Taken together, our data clearly show that NFYA-1 acts cell-autonomously to control PVQ development (Figure 4.3). To reinforce our hypothesis that the PVQ off phenotype was not due to the NFY mutations exclusively affecting the *oyIs14* transgene *(Psra-6::GFP)*, I performed this same analysis for PVQ development for the *nfya-1* mutation crossed with another transgenic strain (*hdIs26)* carrying an independently-generated *Psra-6::RFP* transgene (Figure 4.4).

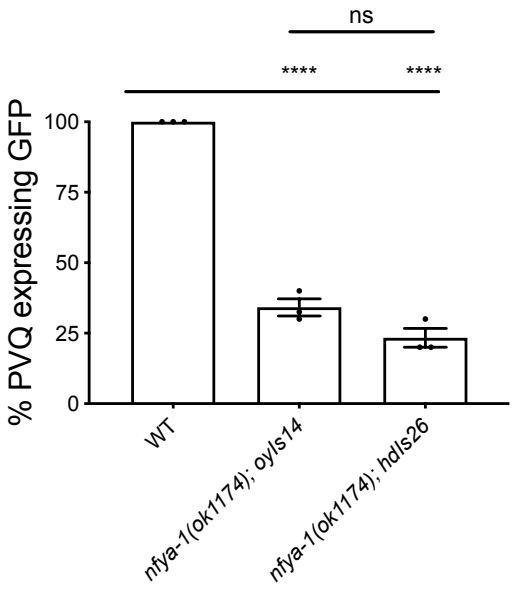

Figure 4.4. Characterization of *nfya-1* mutation for *hdIs26* reporter during PVQ development. The *nfya-1* mutation affects both transgenes (*oyIs14 – Psra-6::GFP* and *hdIs26 – Psra-6::RFP*) in the same manner. These results reinforce the hypothesis that the NFY complex affects PVQ development and not one specific reporter. Data is expressed as mean ± SEM and statistical significance was assessed using multiple comparison ANOVA multiple comparison with Sidak's test. n>45, ns, not significant, ****p<0.0001.

This analysis revealed that both transgenic strains presented around 75% PVQs off in the *nfya-1(ok1174)* mutant. This confirmed that the *nfya-1* mutation is not specifically affecting the oyIs*14* (*Psra-6::GFP*) reporter and that the nfya-1 mutations are causing an impairment in PVQ development. Plus, it is also noticeable that the PVQ neurons in which fluorescence is detectable there is a variation in the intensity in both strains, from very low levels to high (wild-type) levels. This wide range of variability in fluorescence intensity is not observed in wild type animals and lead us to hypothesise that perhaps the NFY complex could be involved in the generation, specification or maintenance of the PVQs.

4.4.2. The role of the NFY complex in PVQ survival

In order to determine if the NFY complex is required for PVQ survival, I examined *nfya-1, nfyb-1* and *nfyc-1* mutant worms for the percentage of PVQs off at two different developmental stages – L1 and L4 (early larval development vs late larval development) (Figure 4.5).

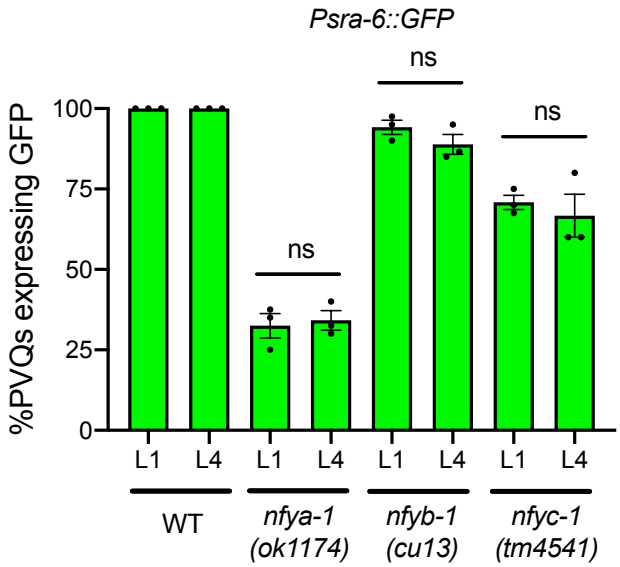

Figure 4.5. Characterization of the role of the NFY complex in PVQ survival. Percentage of *nfya-1, nfyb-1* and *nfyc-1* mutants that do not present GFP in one or both PVQ neurons at L1 stage (early development) and L4 stage (late development). Data is expressed as mean ± SEM and statistical significance was assessed using one-way ANOVA multiple comparison with Tukey's test. n=3 replicates (20 worms per replicate), ns, not significant.

The results obtained do not support the hypothesis that the NFY complex is involved in PVQ maintenance, since both in early and late larval development stages the worms present the same PVQs off penetrance for all the mutant strains analysed.

4.4.3. The role of the NFY complex in PVQ generation

I hypothesized that the NFY complex could be required for the correct generation of the PVQ neurons. The main focus of this analysis was to determine whether the PVQ neurons that did not express any GFP in the previous experiments (Figure 4.2b) are actually being generated. For this purpose, I analysed *nfya-1, nfyb-1* and *nfyc-1* mutant worms crossed with another PVQ reporter strain – *kyIs321* (*Pnpr-11::dsRed*). The results demonstrate that all the NFY complex mutants presented some fluorescence in the PVQs (Figure 4.6). This indicates that the PVQs are present and that the NFY complex is not required for PVQ generation (Figure 4.6).

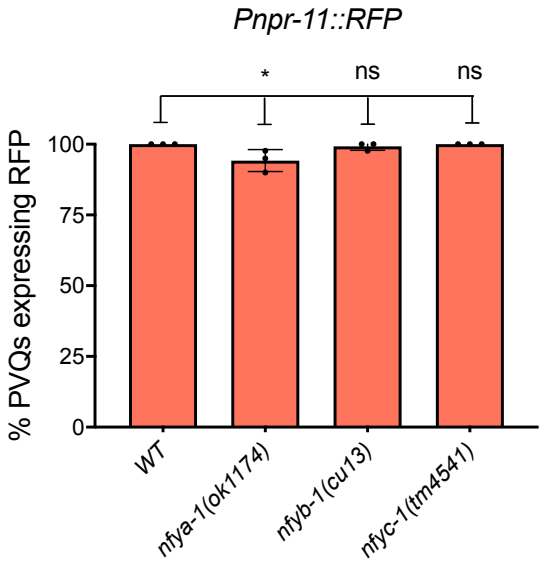

Figure 4.6. The role of the NFY complex in PVQ generation. Mutant worms for the NFY complex were crossed with a RFP PVQ specific reporter (*kyIs321*). The graphic represents the percentage of *nfya-1, nfyb-1* and *nfyc-1* mutants that express RFP in both PVQ neurons. Data is expressed as mean ± SEM and statistical significance was assessed using one-way ANOVA multiple comparison with Dunnet's test. n=3 replicates (20 worms per replicate), ns, not significant, *p<0,05.

Despite the fact that our results do not support the idea that the NFY complex is required for PVQ generation or maintenance, they provide valuable information about a possible role of the NFY complex in PVQ terminal fate acquisition. Since, it is noticeable that mutant worms for the NFY complex present reduced levels of fluorescence when compared with the wild type control throughout development.

Therefore, I measured fluorescence intensity levels (CTCF) in the PVQ to understand if the NFY complex mutations could be affecting specific traits of the PVQ identity.

4.4.4. NFY complex controls the expression of a repertoire of PVQ terminal genes

To determine the extent that the NFY complex is essential for PVQ terminal fate acquisition, I measured the fluorescence levels (Corrected Total Cell Fluorescence CTCF) in the PVQs of *nfya-1, nfyb-1* and *nfyc-1* mutant worms crossed with the GFP and RFP reporter strains used previously in this work (*oyIs14* and *kyIs321*, respectively) (Figure 4.7a & b).

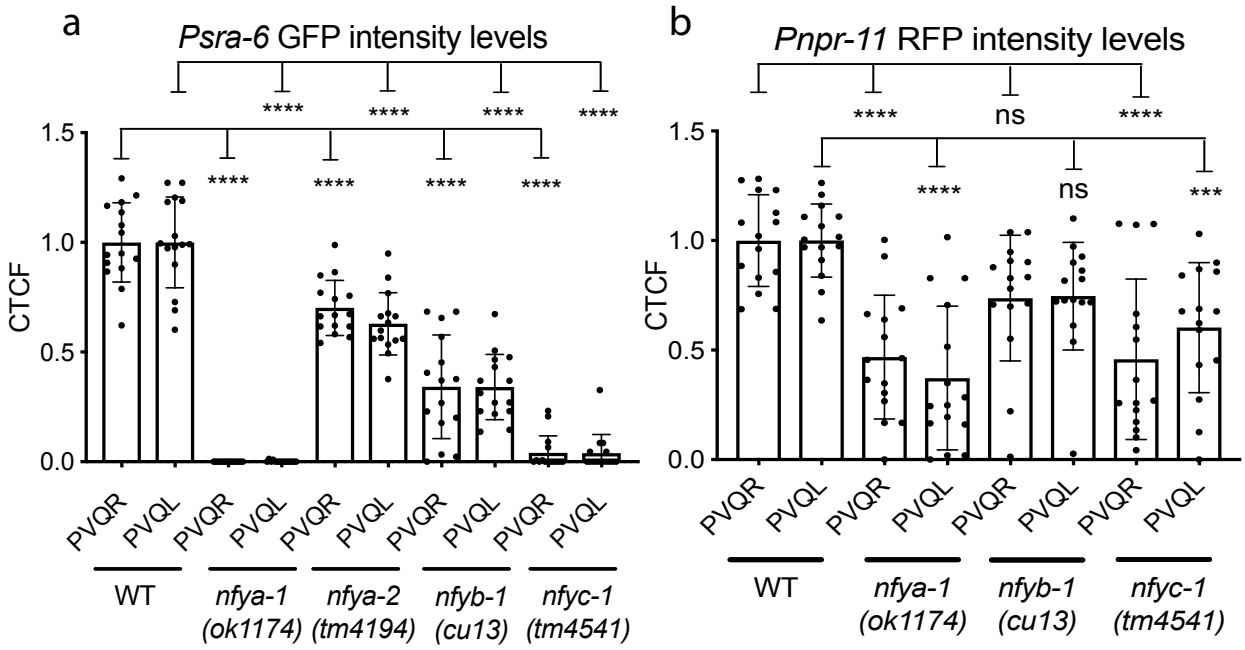

Figure 4.7. Intensity of GFP/RFP expression (CTCF) of *nfya-1, nfya-2, nfyb-1* and *nfyc-1* mutants normalized against wild type. The results are divided into PVQR or PVQL for all the mutants. **A)** – Quantification of the GFP levels for the NFY complex mutant worms crossed with the *oyIs14* reporter. **B)** – Quantification of the RFP levels for the NFY complex mutant worms crossed with the *kyIs321* reporter. Data is expressed as mean ± SD and statistical significance was assessed using one-way ANOVA multiple comparison with Sidak's test. n=15 neurons, ns, not significant, *p<0,05, **p<0,01, ****p<0,0001.

The results show a severe reduction in the fluorescence levels in both markers for all the mutants strains with the exception of the *nfya-2* mutants that present a very mild phenotype. Moreover, *nfya-1* allele presents the most severe phenotype, with 100% reduction and almost 50% for the RFP and GFP intensity, respectively. These results reflect the NFY complex's importance in correct PVQ specification, and reinforces results from Figure 4.2b.

To further examine the NFY complex role in PVQ specification I analysed 12 additional PVQ terminal fate and pan-neuronal reporters – *mir-124 (MIcro RNA 124), srh-277 (Serpentine Receptor, class H), sri-1* (*Serpentine Receptor, class I), srg-32 (Serpentine Receptor, class G (gamma)), srv*-32 (Serpentine Receptor, class V)*, egl-47A* and *B (G protein-coupled receptor)*, *dop-1 (DOPamine receptor)*, *rab-3 (Rab small GTPases), ceh-43 (*Homeobox TF) and *zig-5 (2 (Zwei) IG domain protein)* (Figure 4.8a, b & c). These reporters represent different aspects of PVQ physiology and molecular structure- All the reporters analysed have at least one CCAAT binding site in their promoter region, suggesting regulation by the NFY complex (Figure 4.8c).

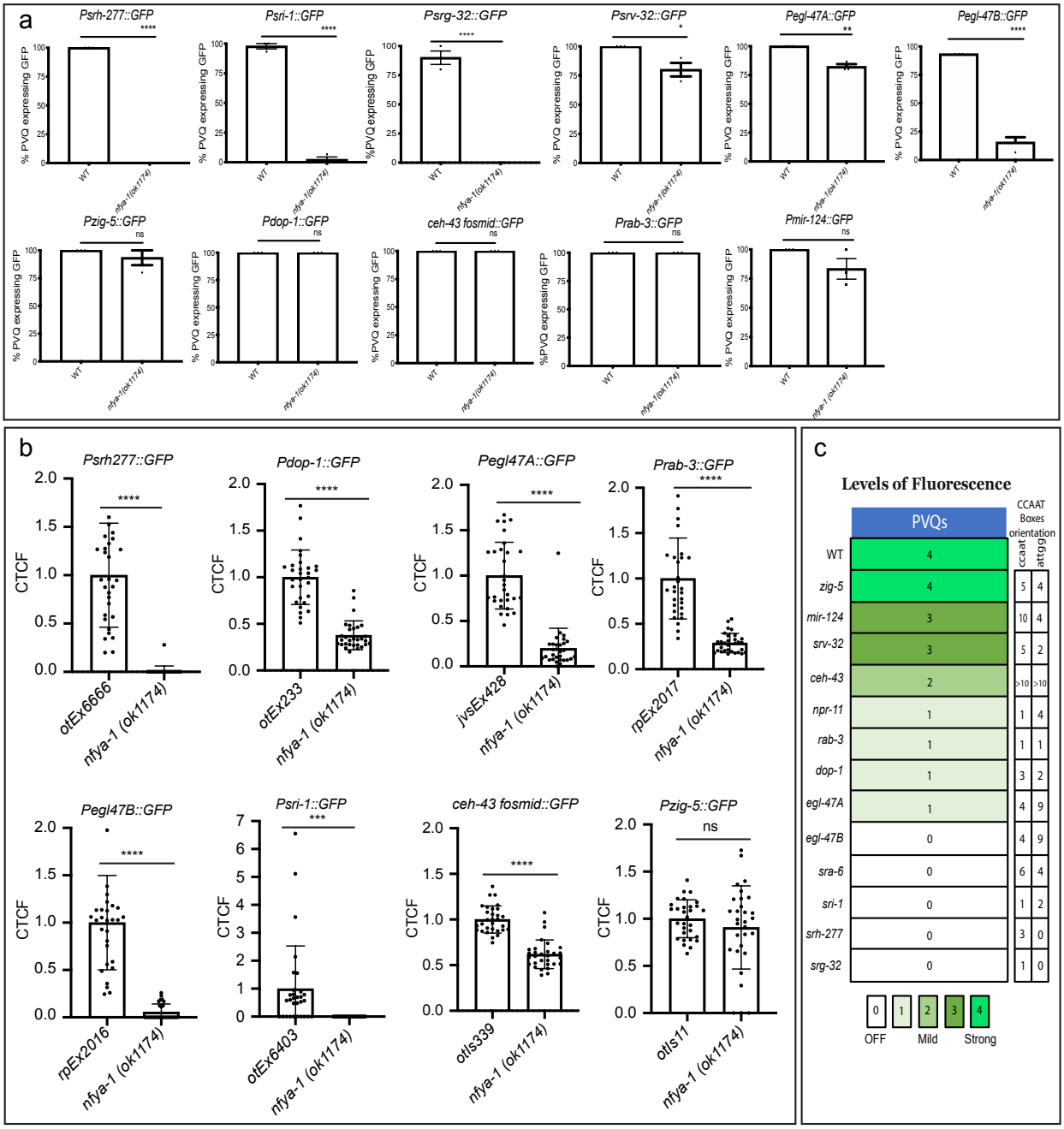

Figure 4.8. Characterization of *nfya-1* gene function during PVQ terminal fate acquisition. A - Mutant worms for the NFYA-1 subunit were crossed with different transgenic reporters (*Psrh-277::GFP; Psri-1::GFP; Psrg-32::GFP, Psrv-32::GFP, Pegl-47A::GFP, Pegl-47A::GFP, Pzig-5::GFP, Pdop-1::GFP, ceh-43 fosmid::GFP, Prab-3::GFP* and *Pmir-124::GFP;*) and were scored for PVQ off. The data represent the percentage of *nfya-1* mutant animals that present fluorescence in one or both PVQs. **B -** GFP intensity was measured for the following reporters, *Psrh-277::GFP; Psri-1::GFP; Psrv-32::GFP, Pegl-47A::GFP* and *Pegl-47A::GFP, ceh-43 fosmid::GFP*, *Pzig-5::GFP* and *Prab-3::GFP* in wild type and *nfya-1* mutant background. **C –** The heat map represents all the markers analysed in this work. The heat map demonstrates that some of the markers are severely affected by the *nfya-1* mutation (white boxes with a 0), while other markers are mildly or very mildly affected by the *nfya-1* mutation (light green boxes with a 2 or a 3). Also, we investigate if the reporters used in this work have CCAAT box in the promoter

region. Data is expressed as mean ± SEM with n=3 replicates (>45 worms per replicate) for A and C. Data is expressed in mean ± SD with n=30 neurons for D. Statistical significance was assessed using t test., ns, not significant, **p<0,005, ***p<0,001, ****p<0,0001.

This analysis showed that fluorescence of the PVQ reporters used in figure 4.8a & b is reduced in *nfya-1* mutant worms when compared to wild type. Both *srh-277, sri-1 and srg-32* reporters are severely affected by the *nfya-1* mutation, while the *srv-32* and *mir-124* reporters presented a much milder phenotype in terms of PVQs defects. In the case of *mir-124*, the data shows a mild downregulation in GFP expression and a mild percentage of PVQs off. Also, other reporters for the *egl-47* isoform A and B, *rab-3, ceh-43*, and *dop-1* receptor present a milder phenotype than *srh-277* and *sri-1,* but more severe than *srv-32* and *mir-124* (Figure 4.8a). Lastly, we observed that the zig-5 reporter is not affected and exhibits the same levels seen in the control worms (Figure 4.8b).

Taken together, these results (summarized in Figure 4.8c) show that the NFY complex plays an essential role during PVQ fate acquisition by regulating a gene battery required for PVQ fate.

4.4.5. The NFY complex directly regulates PVQ gene expression

To elucidate whether the NFY complex directly regulates gene expression in the PVQ neurons, I selected the most affected reporters by loss of *nfya-1*, *srh-277* and *sri-1* for further analysis. I mutagenized the CCAAT binding site that was closest to the translational start site (Figure 4.9). Subsequently, I quantified the percentage of worms that showed PVQs-off in comparison to the original reporter strain (Figure 4.9a & b).

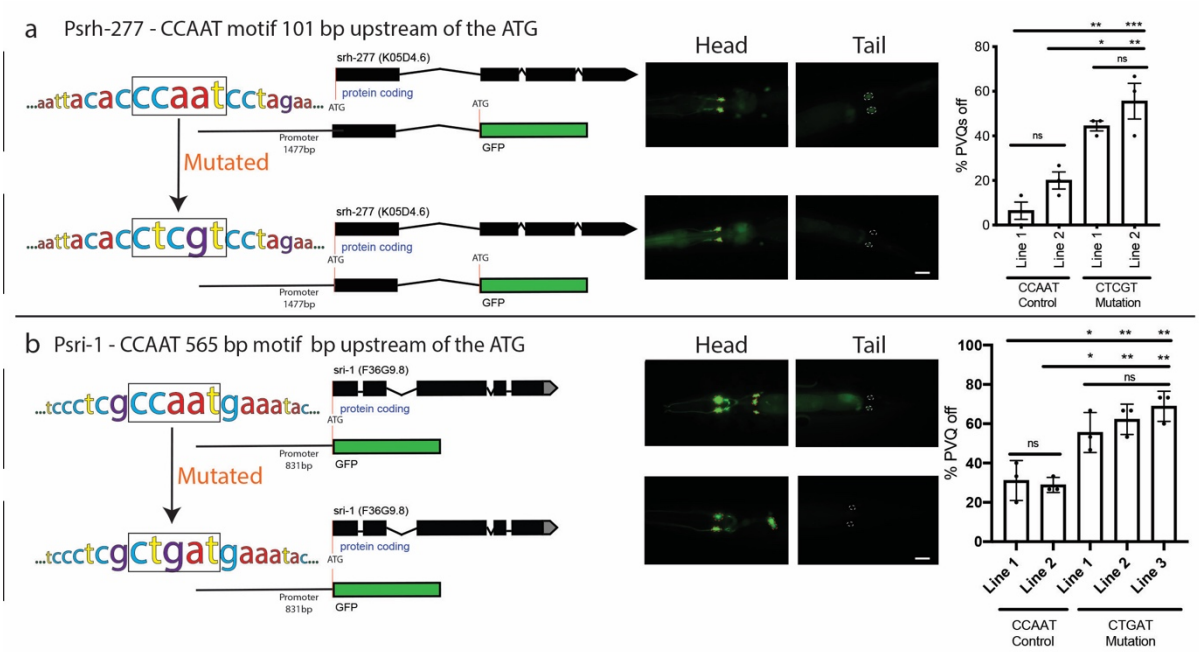

Figure 4.9. Mutagenesis analysis of *sri-1* and *srh-277* promoter regions. A - Mutagenesis analysis performed on the CCAAT site closest to the *srh-277* translational start site (exon 1). The wild type promoter sequence (with the CCAAT site highlighted) and the mutagenized promoter sequence (with the CTCGT highlighted) are shown on the left. The GFP coding sequence was inserted after the first exon coding sequence. Representative images of GFP expression in the head and tail for both transgenic strains injected with the wild type or mutagenized sequences are seen on the right. The graph represents the GFP intensity measurements in the PVQs (white circles) for control and mutagenized strains. **B -** Mutagenesis analysis performed on the CCAAT site closer to the *sri-1* transcription start site (exon 1). The wild type promoter sequence (with the CCAAT site highlighted) and the mutagenized promoter sequenced (with the CTGAT site highlighted) are represented in the left. The GFP coding sequence was inserted after the promoter sequence. Representative images of GFP expression in the head and tail for both transgenic strains injected with the wild type (top) and mutagenized sequences (bottom) are seen on the right. The graph represents the GFP intensity measurements in the PVQs (white circles) for control and mutagenized strains. Data is expressed as mean ± SD for 2-3 independent lines precondition and statistical significance was assessed using a one-way ANOVA multiple comparison with Sidak's multiple comparison test. n=3 replicates (15 worms per replicate), ns, not significant, *p<0,05, **p<0,005, ***p<0,001. Scale bar (10 µm).

These results show that mutating the CCAAT NFY binding site in the *srh-277* and *sri-1* promoters causes a reduction of reporter expression only in PVQs (ON to OFF) and suggest that the NFY complex directly regulates the transcription of these two genes specifically in the PVQs.

4.4.6. The NFY complex controls pan-neuronal gene expression in motor neuron

Bioinformatics analysis performed on ChIP-Seq data from the ENCODE project (modencode.org) for the NFY complex revealed binding peaks in multiple pan-neuronal genes, such as *rab-3, unc-11* and ric-4. To determine if the NFY complex has a broader function in the *C. elegans* nervous system, I analysed the expression levels of *rab-3* pan-neuronal reporter. *rab-3* encodes a small GTPase that regulates exocytosis during synaptic transmission. I found that *rab-3* expression is reduced in the PVQs of *nfya-1* mutants. This suggests that in addition to PVQ-specific traits, the NFY complex controls the expression of genes that control general neuronal function.

I next broadened my analysis to determine the function of NFY in other neurons. The analysis was carried out in five motor neurons located within the ventral nerve cord. First, I found that both the *nfya-1* and the *rab-3* reporters are expressed in these neurons (Figure 4.10a). Moreover, *nfya-1* mutant worms have a 20% decrease in rab-3::GFP intensity in the motor neurons when compared with wild type worms (Figure 4.10b). This result suggests that the NFY complex plays a role in controlling gene expression of other neuronal cell types beyond the PVQs. To confirm this result, I analysed the expression of two additional pan-neuronal reporters, *unc-11 and ric-4.* The results showed that the *unc-11* reporter is 25% downregulated in the *nfya-1* mutant worms (Figure 4.10c). In contrast, the *ric-4* fosmid did not show any significant difference between wild type and mutant worms (Figure 4.10c). These results suggest that the NFY complex is required for the regulation of pan-neuronal gene expression. However, since the *ric-4* fosmid was not affected, it seems that other regulatory elements are required to regulate this gene.

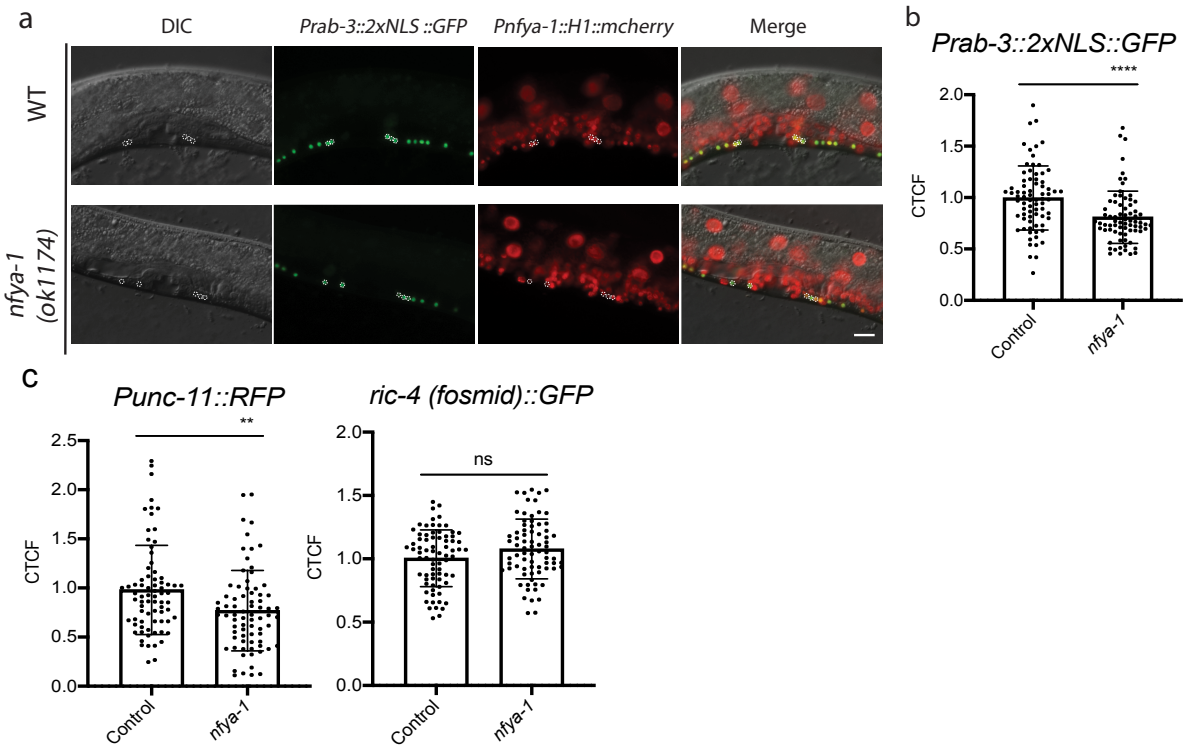

Figure 4.10. Characterization of the *nfya-1* role in motor neurons development. Intensity of GFP expression (CTCF) in the *nfya-1* mutant normalized against wild type worms. **A –** Co-localization experiments for the NFYA-1 subunit (red) and *Prab-3::2xNLS::GFP* were performed at the L4 stage. These experiments revealed that the NFYA-1 subunit is expressed in the nuclei of the motor neurons surrounding the vulva at L4 stage due to its co-localization with the *rab-3* reporter. **B –** GFP intensity measured in the neurons marked with white circles (DB5, AS6, VD7, DA5 and VC5). **C –** GFP intensity measured in the neurons marked as in A for *Punc-11::GFP* and *ric-4 (fosmid)::GFP reporters*. Data is expressed as mean ± SD and statistical significance was assessed using an unpaired t-test. n=75 neurons, ns, not significant, **p<0,005, ****p<0,0001. Scale bar (10 µm).

4.4.7. NFY mutants display VNC axon guidance defects

Here, I investigated if the NFY complex plays a role in other steps of neuronal development and ventral nerve cord formation (VNC). For this analysis, I analysed axon guidance in the different *nfya* (1 and 2), *b* and *c* mutant worms (Figure 4.11a – arrow head).

Interestingly, *nfya-1, nfyb-1* and *nfyc-1* mutants present ~30% penetrant PVQ axon guidance defects (Figure 4.11b). Conversely, *nfya-2* mutant worms do not show any defects when compared with wild type worms (Figure 4.11b). These results suggest

that the NFYA-2 subunit may not be involved in PVQ axon guidance. Reiterating that the nfya-1 and nfya-2 subunits may possess independent roles.

To confirm that the defects observed were a consequence of mutation in the *nfy* genes, I performed a rescue experiment to revert the axon guidance phenotype. For this analysis, I re-supplied the *nfya-1* mutant worms with a wild type copy of the *nfya-1* gene, and I was able to fully rescue the *nfya-1* axon guidance phenotype. These data reveal for the first time that the NFY complex is required for the correct axon guidance (Figure 4.11b).

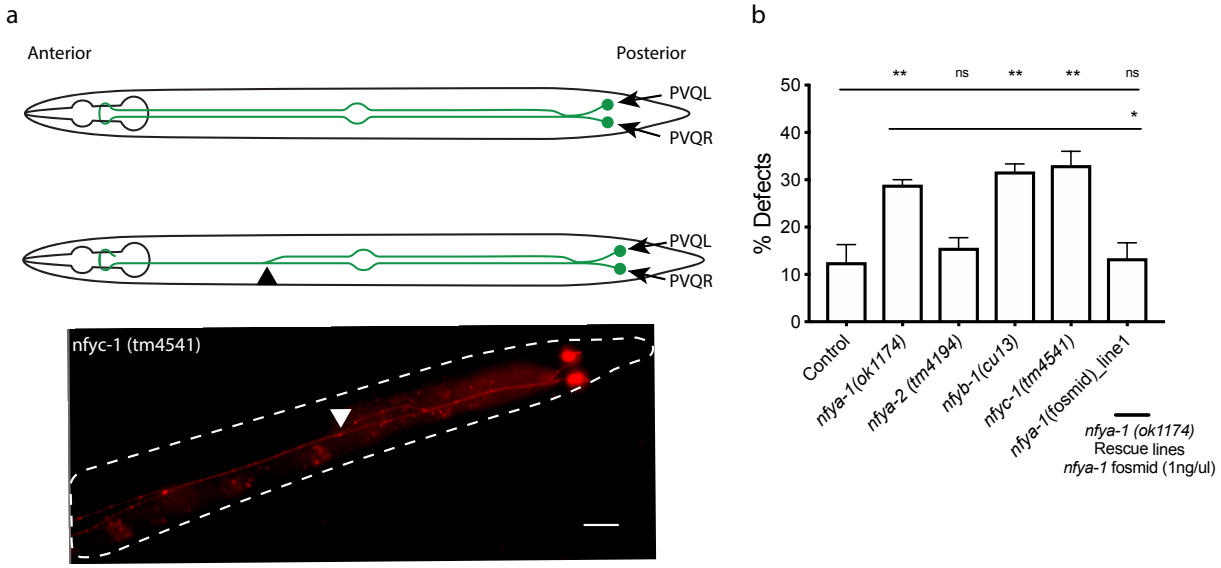

Figure 4.11. Characterization of *nfya-1, nfyb-1 and nfyc-1* mutations in PVQ axon development. Mutant worms for each *nfy* gene were crossed with a reporter strain that expresses RFP in the PVQ neurons, and scored for PVQ developmental defects. A – Schematic representation of a wild type worm with normal PVQs on the top and a NFY mutant worm with a guidance defect on the bottom (black arrow head points to the region where the PVQL crosses the midline and merges with the PVQR). **Bottom** –*nfyc-1* mutant worm presenting a guidance defect (white arrow head points to the region where the PVQL crosses the midline and merges with the PVQR)**. B -** The *nfya-1*, *nfyb-1* and *nfyc-1* mutants present a very similar phenotype with around 30% of axon guidance defects when compared with the control strain (10%). The majority of *nfya-1* and *nfyc-1* mutant worms (around 60%) could not be scored due to the reduced levels of fluorescence in the PVQs. Nevertheless, approximately 30% of the remaining animals present axon guidance defects. On the opposite side, *nfya-2* mutant worms do not present any difference in guidance defects when compared to wild type worms. The fosmid rescue line is able to fully rescue the guidance defect to wild type level. Data is expressed as mean ± SEM and statistical significance was assessed using one-way ANOVA multiple comparison with Tukey's test. n=3 replicates (>15 worms per replicate), ns, not significant, *p<0,05, **p<0,005. Scale bar (10 µm).

4.5. Discussion

In this work, I identified the NFY complex as a novel regulator of the development of a pair of glutamatergic interneurons (PVQs). Moreover, I characterized the role of the NFY complex in the PVQs, and the function of the different NFY subunits during neuronal differentiation.

Previous studies have shown that conditional suppression of the NFYA subunit leads to the neurodegeneration of specific post-mitotic neurons in mice (Yamanaka *et al.*, 2014; Yamanaka *et al.*, 2016). However, my work shows that the NFY complex is not required for PVQ generation or survival (Figure 4.5 & 4.6). This suggest that the NFY complex may play specific roles in different neuronal subtypes. This hypothesis is supported by the fact that inactivation of the NFY complex in different neurons in the central nervous system of mice results in specific neuronal pathologies (Yamanaka *et al.*, 2014; Yamanaka *et al.*, 2016).

To examine whether the NFY complex is required for PVQ terminal fate acquisition, I analysed different reporters expressed in the PVQs. I found that the NFY complex is essential to regulate the expression of most PVQ terminal fate reporters. Interestingly, the NFY complex is crucial for the expression of specific PVQ neuronal fate reporters, but plays a minor role in regulating other PVQ-expressed genes. For example, analysis of the *nfya-1* mutant showed that *sra-6, srh-277* and *sri-1* reporters are highly downregulated (Figure 4.8). In contrast, both *srv-32* and *mir-124* reporters are only mildly affected, and the *zig-5* reporter is not affected (Figure 4.8). Moreover, mutagenesis analysis suggested that the NFY complex directly regulates the expression of the *sri-1* and *srh-277* genes. However, the mutagenesis analysis only revealed a partial increase in the percentage of PVQs off. This could be explained by the fact that the promoters of both genes have more than one CCAAT bind motif, which contribute to regulation of expression by the NFY complex.

Based on these findings, I hypothesize that the NFY complex could work alone and/or in concert with other transcription factors to directly bind to different sites within the

promoter region of these terminal selector genes responsible for PVQ fate. This hypothesis is in line with previous work done by Oldfield *et al.* (2014) that revealed that the NFY complex promotes chromatin accessibility for other cell specific transcription factors to perform their function.

Notably, knockouts of each NFY complex subunit (A, B and C) show different levels of regulation for the same marker. As an example, the *sra-6* reporter shows approximately 70% PVQ-off in the *nfya-1* mutant, while *nfyb-1* and *nfyc-1* mutants exhibit 20% and 40% PVQs-off, respectively. These differences are possibly due to specific roles that each subunit plays within the NFY complex. Based on previous interaction studies conducted by Deng *et al*. (2007), it is possible that the NFY complex interacts with different co-activators to control the expression of different genes. Also, Hu and colleagues previously described that the NFYC and NFYA subunits possess a poly-glutamine region essential for the NFY complex function (Hu and Maity, 2000). This could be another possible explanation as to why both the *nfyc-1* and *nfya-1* mutant worms tend to show a more severe phenotype compared with the *nfyb-1* mutant worms.

Another interesting point to be drawn from my results is that *nfya-2* mutant worms do not present PVQ axon guidance defects (Figure 15a). This result reinforces the hypothesis tested by Deng and colleagues (2007) in which both NFYA-1 and 2 are able to form complexes with the NFYB/NFYC dimer, and therefore may have different functions. To support this idea, my data shows that *nfya-1* mutants present guidance defects while *nfya-2* mutants do not.

Previous studies revealed that the NFY complex is required for the development of different neuronal types in mammals (Milton *et al.*, 2013; Yamanaka *et al.*, 2016). Moreover, ChIP-Seq data revealed putative binding sites in *C. elegans* pan-neuronal genes, such as *unc-11*, *rab-3* and *ric-4*. Therefore, I hypothesized that the NFY complex also regulates pan-neuronal gene expression. I focused my analysis on the expression of *rab-3*, *unc-11* and *ric-4* markers in a group of five motor neurons that surround the vulva. My results showed downregulation (25 to 30%) of *rab-3* and *unc-*

11 expression in the *nfya-1* mutant when compared with wild type worms (Figure 4.10). Interestingly, *ric-4* reporter expression levels were not affected in *nfya-1* mutant worms, suggesting that there are other regulatory elements in the *ric-4* reporter responsible for maintaining correct expression pattern in addition to the NFY complex motif. These results reinforce the hypothesis that the NFY complex is not only required for specific neuronal fate acquisition (such as PVQ fate) but also acts as a master regulator of neuronal development, and possibly neuronal function by regulating pan-neuronal features, such as synaptic proteins, in different neuronal subtypes. In my final discussion, I will continue to explore the topic of pan-neuronal gene regulation.

Finally, I found that the NFY complex is also required for correct PVQ axon guidance during development. This result is in line with recent RNA-Seq data that shows that the NFY complex could be involved in the regulation of genes required for axon guidance, such as *sax-2*, *unc-71*, *unc-104*, further supporting the importance of the NFY complex for PVQ development (Tharyan *et al.*, 2020).

4.6. Future directions:

- In chapter 5, I genetically engineered the NFY complex subunits using CRISPR/Cas9 technology to study the interaction between the NFY complex and other proteins during neuronal fate acquisition.
- Finalise characterizing the NFY complex function in PVQ development. I aim to perform further analysis in the *nfy* mutant worms to investigate the role of the NFY complex in PVQ development. I aim to publish this work in 2020.

5. Endogenously tagging the NFY complex subunits using CRISPR-Cas9 to enable proteomic analysis.

5.1. Abstract

In chapter 4, I demonstrated that the NFY complex controls the expression of PVQ-specific and general genes that control neuronal function. The NFY complex is known to directly regulate gene expression through interactions with other proteins, such as p300, MES-2/MES-6, and ATF6α. However, NFY complex interactors in *C. elegans* are undescribed, as is how these interactions drive neuronal development.

In this chapter, I developed molecular tools to allow dissection of the molecular mechanisms that control NFY-mediated nervous system formation. Using CRISPR-Cas9 technology, I generated functional endogenously tagged NFYB-1 and NFYC-1 subunits by adding an HA-tag to the N-terminus of each protein. Additionally, I designed a V5-tagged NFYA-1 strain (N-terminally tagged) that was generated by SunyBiotech. The V5::NFYA-1, HA::NFYB-1 and HA::NFYC-1 subunits were functionally validated using two independent assays and will be used for further characterization of the NFY complex function in the future.

In this chapter, I have detected HA::NFYB-1 (46 kDa), the V5::NFYA-1 (53.6 kDa) using western blot, and successfully performed immunoprecipitation of the HA::NFYB-1 (46 kDa).

5.3. Introduction

Transcriptional regulation is an essential mechanism for controlling gene expression. In eukaryotes, it is thought that most of TFs recruit co-activators/repressors to assemble a transcriptional complex to regulate gene expression (Reiter *et al.*, 2017). As an example, the NFY complex needs to interact with specific proteins to regulate gene expression. Previous *in vitro* studies performed by Deng *et al.* (2007) showed that the NFY complex is required to interact with the polycomb Group Proteins (PcG) MES-2/MES-6 to repress *egl-5* expression and thus regulate tail patterning in *C. elegans*. Additionally, Li *et al.* (1998) presented evidence that the NFY complex interacts with the p300 co-activator/acetyltransferase to disrupt chromatin conformation, and thus enhance gene expression of specific genes, such as *hsp70*. However, the full repertoire of NFY complex interactions and how this drive neuronal development remains uncharacterised.

In chapter 4, I demonstrated that the NFY complex is required to regulate expression of certain neuronal genes, likely through direct binding. However, the molecular mechanisms required for the NFY complex perform its function are unclear.

Here, I aimed to genetically label the NFY complex subunits (NFYA, NFYB, and NFYC) to identify novel co-activators required for the NFY complex function in neuronal development through mass spectrometry.

5.2. Methods

Strains

All strains were maintained at 20°C for two generations prior to scoring, as previously described (Brenner, 1974), unless stated otherwise. All the transgenic and mutant strains used in this work are detailed in Table 2.1 and Supplemental Table 3.

Genetics

The following mutant strains were used in this work: *nfya-1(ok1174), nfyb-1(cu13),* and *nfyc-1(tm4145)* (Supplemental Table 2). The mutant strains represent strong loss-of-function alleles as most of the protein sequence is deleted. The CRISPR/Cas9 manipulation of the worms follows the protocol explained in the methodology section 2.6 to create the following strains – *ROP01(*syb2237), *RJP4355(rp132), RJP4353(rp130)* (Supplemental Table 3). CRISPR strains were crossed with *oyIs14* and outcrossed with WT to generate the strains for analysis in this chapter - *RJP4645, RJP4427, RJP4496, RJP4428, RJP4497* (Supplemental Table 3). The protocols regarding injection and CRISPR-Cas9 methodology can be found in sections 2.3 and 2.6.

Brood size analysis

Brood size analysis was performed as described in section 2.4.

Expression constructs and transgenic animals

The reporter strains used in this study are described in Supplemental Table 3.

Fluorescence microscopy

GFP expression and neuron morphology analyses were performed in L4 hermaphrodites anesthetized in 10 mM sodium azide (NaN_3) and mounted on 5% agarose on glass slides. Images were capture using an automated fluorescence microscope (Zeiss, AXIO) and ZEN software. Fluorescence levels were measured using FIJI software (ImageJ V.1.47n). CTCFs measurements were performed as described in section 2.9.

Proteomics

Refer to section 2.7 for the protocol on protein extraction, western blotting, and immunoprecipitation.

Statistical analysis

Statistical analyses and graphs were generated in GraphPad Prism 7 (www.graphpad.com). The statistical analysis and test used are indicated in the respective figure legend. Differences with a *p value<0.05 were considered significant.

5.4. Results

5.4.1. Endogenously tagging NFYA, B and C subunits using CRISPR technology

To study how the NFY complex regulates gene expression, I wished to design tools to enable proteomics and ChIP-PCR/ChIP-Seq analyses. These analyses will identify interactors of the NFY complex, and identify candidate genes that are regulated by the NFY complex. To this end, I generated epitope-tagged CRISPR-generated worms for each NFY subunit (Figure 5.1).

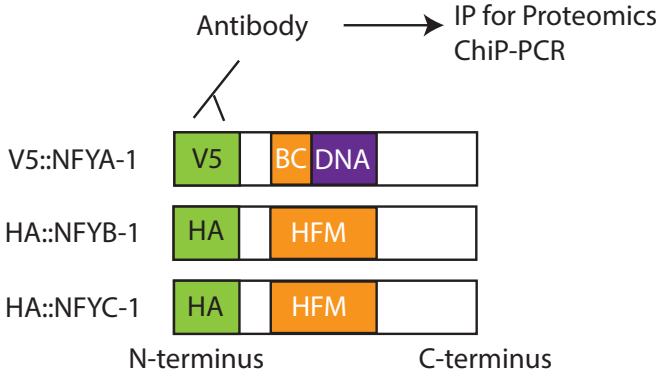

Figure 5.1. Schematic representation of the endogenously tagged NFY subunits designed in this work (HA or V5 tag) at the N-terminus using CRISPR-Cas9 technology. Histone folding and BC motifs in orange and DNA binding region in purple.

To date, tagging of the NFY complex has been carried out in cell culture for the NFYC-1 subunit and tagging of the NFYA-1 subunit in *C. elegans*, by the ENCODE project (encodeproject.org). To tag the NFY subunits, I used CRISPR/Cas9 genome editing (Dokshin *et al.* (2018). I started by tagging the NFYA-1 subunit with GFP (870 bp), but I was never able to detect the GFP insertion by PCR. This possibly was due to the large size of the GFP sequence and/or low insertion efficiency. Then, I tried to tag the NFYB-1 subunit with a 3xFLAG-tag (75 bp) as its smaller size (compared to GFP) would make the insertion more feasible. Although the 3xFLAG-tag was inserted in the correct position, sequencing analysis revealed that the 3' region of the gene promoter was not repaired correctly and possibly a second 3xFLAG-tag was also inserted in this region (Figure 5.2). Accordingly, homozygous worms for the FLAG-tag insertion presented a

phenotype similar to the *nfyb-1* mutant phenotype with a small brood size (described in section 5.4.2), confirming that the NFYB-1 subunit was non-functional (Figure 5.2).

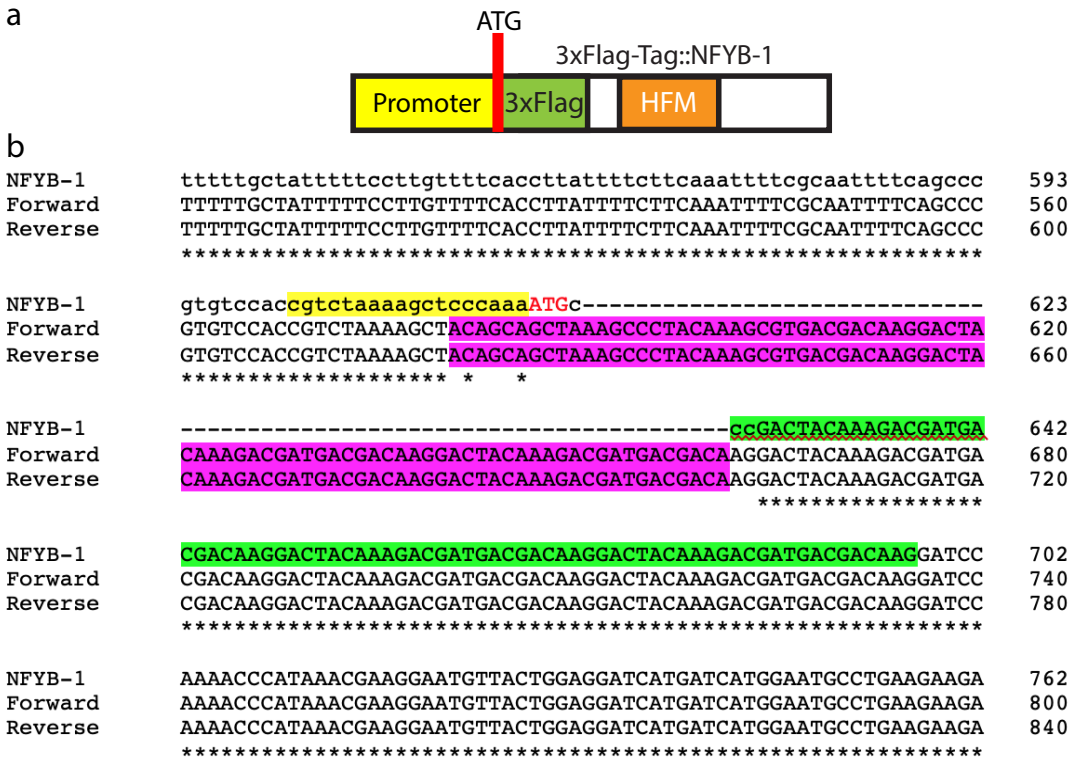

Figure 5.2. Sequencing results for the CRISPR-Cas9 3xFLAG-tag of the NFYB-1 subunit. A - Schematic representation of the NFYB-1 subunits tagged with 3xFlag-tag (N-terminus). **B -** The upper sequence shows the predicted sequence for tagging the *nfyb-1* gene with 3xFLAG-tag (75 bp). Forward and reverse sequences refer to Sanger sequencing results obtained from the experiment. Highlighted in yellow is the 3' region of the nfyb-1 promoter and highlighted in purple is the portion of DNA incorrectly repaired. It seems that a second 3xFLAG fragment was inserted, and the repair process occurred incorrectly. Highlighted in green is the 3xFLAG-tag sequence and labelled in red is the start codon of the *nfyb-1* gene. (*) – represent similarities in the alignment sequence. Capital letters in the *nfyb-1* sequence represent the coding region of the sequence.

To improve the efficiency of the insertion and reduce the possibility of affecting the function of the NFY complex, we next decided to use a smaller tag (HA-tag encoded by 27 bp). I successfully inserted the HA-tag in-frame at the N-terminus of the NFYB-1 and NFYC-1 subunits as depicted in the sequencing analysis presented below (Figure 5.3 a & b).

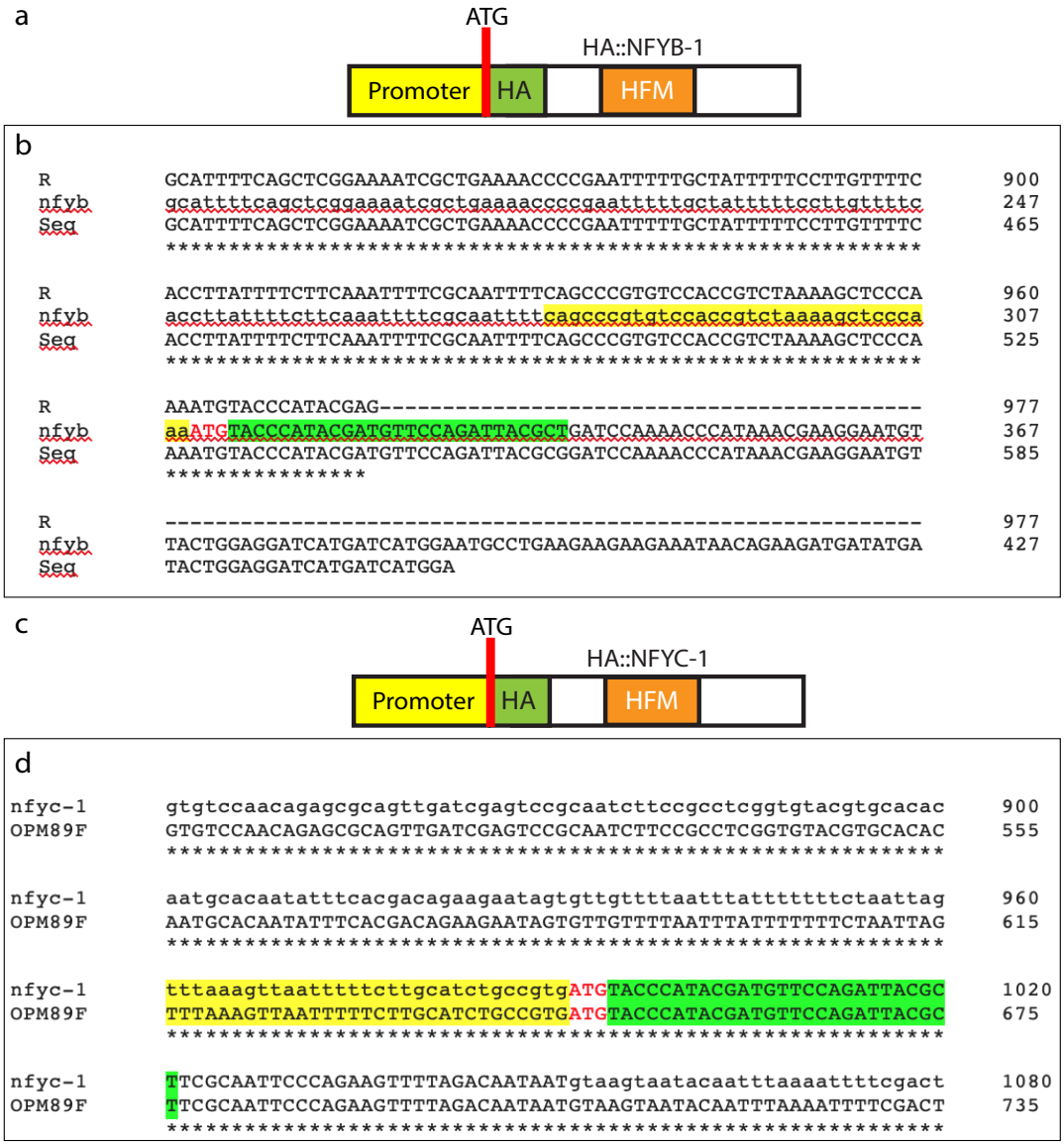

Figure 5.3. Sequencing results for the CRISPR-Cas9 HA-tag of the NFYB-1 and NFYC-1 subunits. **A -** Schematic representation of the NFYB-1 subunits tagged with HA-tag (N-terminus). **B -** The top sequence (R sequence) refers to the DNA obtained after sequencing the DNA (with a reverse primer) from worms that were positive for HA-tag insertion (27 bp). The middle sequence represents the reference DNA template for *nfyb-1* gene with HA-tag. The bottom sequence (Seq) refers to the DNA obtained after sequencing the DNA (with a forward primer) from transgenic worms that were positive for HA-tag insertion. The 3' region of the *nfyb-1* promoter is highlighted in yellow. C - Schematic representation of the NFYC-1 subunits tagged with HA-tag (N-terminus). **D -** The top sequence represents the reference DNA template for *nfyc-1* gene with HA-tag. The OPM89F sequence refers to the DNA obtained from worms that were positive for HA-tag insertion. The 3' region of the *nfyc-1* promoter is highlighted in yellow. The HA-tag sequence is highlighted in green. The start codon of the *nfyc-1* gene is labelled in red. (*) – represent similarities in the alignment sequence. Capital letters in the *nfyc-1* sequence represent the coding region of the sequence.

Finally, I designed a strain with the NFYA-1 protein tagged with V5-tag (42 bp) in the N-terminal region and ordered it from SunyBiotech. The sequencing result presented by the SunyBiotech company is illustrated in Figure 5.4.

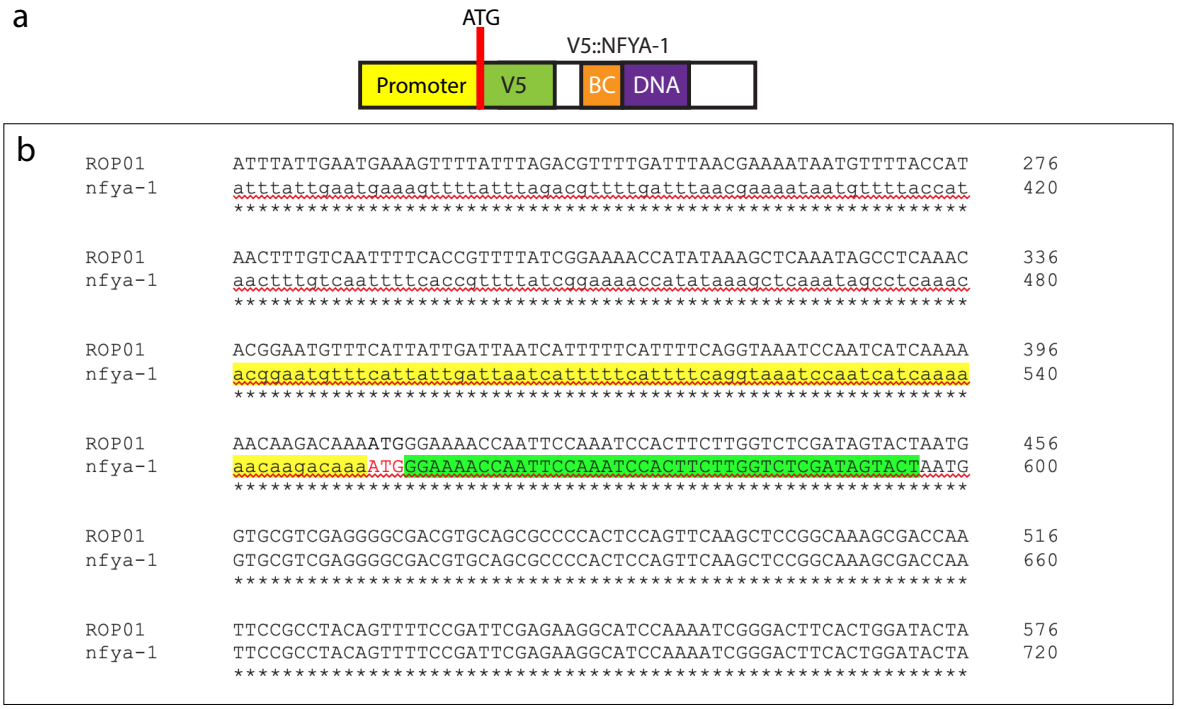

Figure 5.4. Sequencing results for the CRISPR-Cas9 V5-tag of the NFYA-1 subunit. A - Schematic representation of the NFYB-1 subunits tagged with HA-tag (N-terminus). **B -** The top sequence (R sequence) refers to the DNA obtained after sequencing the DNA from transgenic worms that were positive for V5-tag insertion (42 bp). The bottom sequence represents the reference DNA template for the *nfya-1* gene with V5-tag. The 3' region of the *nfya-1* promoter is labelled in yellow, the HA-tag sequence is labelled in green, and the start codon of the *nfya-1* gene is labelled in red. (*) – represent similarities in the alignment sequence.

5.4.2. Functional validation of endogenously tagged NFY complex subunits

To examine whether the CRISPR-Cas9 tagged NFY subunit strains were functional, I analysed the brood size and percentage of PVQs-off (without GFP expression). The brood size, i.e. number of progeny produced by a single worm throughout its lifespan, was based on previous studies performed in the NFY mutant worms that reported low progeny number when compared to wild type (Deng *et al.*, 2007). To perform this

analysis, the number of embryos/worms that a single L4 hermaphrodite worm produced over a period of 6 days was recorded (Figure 5.5). The results from this experiment clearly show that HA::NFYB-1 tagged worms generate the same number of progeny when compared with the wild type (300 embryos). This contrasts to the low brood size of *nfyb-1* mutant worms (Figure 5.5a). The same result was observed for HA::NFYC-1 worms that generate the same number of progeny as wild type worms (300 embryos), in contrast to *nfyc-1* mutant worms (Figure 5.5b). Interestingly, mutant worms for both *nfyb-1* and *nfyc-1* show a lower number of unfertilized eggs when compared with the wild type and HA::NFYB-1 and HA::NFYC-1, respectively. Since the number of progeny is reduced and the number of fertilized eggs is also reduced, these observations point to an underlying problem relating with oocyte development, e.g. low number of oocytes, and appear to exclude any problems related with sperm development or fertilization. Further, I used the same validation assays for the V5::NFYA-1 CRISPR-Cas9 strain. The results showed that V5::NFYA-1 transgenic worms present the same brood size as wild type worms (300 embryos), again in contrast to the low brood of *nfya-1* mutant worms (Figure 5.5c).

For analysis of PVQ cell fate specification, the CRISPR-tagged strains were crossed into a stable PVQ transgenic reporter - *Psra-6::GFP (oyIs14)*. I found that V5::NFYA-1, HA::NFYB-1 and HA::NFYC-1 worms exhibited wild-type expression of *Psra-6::*GFP in the PVQs (Figure 5.6a). This contrasts the *nfya-1*, *nfyb-1* and *nfyc-1* mutant worms that exhibit loss of PVQ expression (Chapter 4, Figure 4.2b).

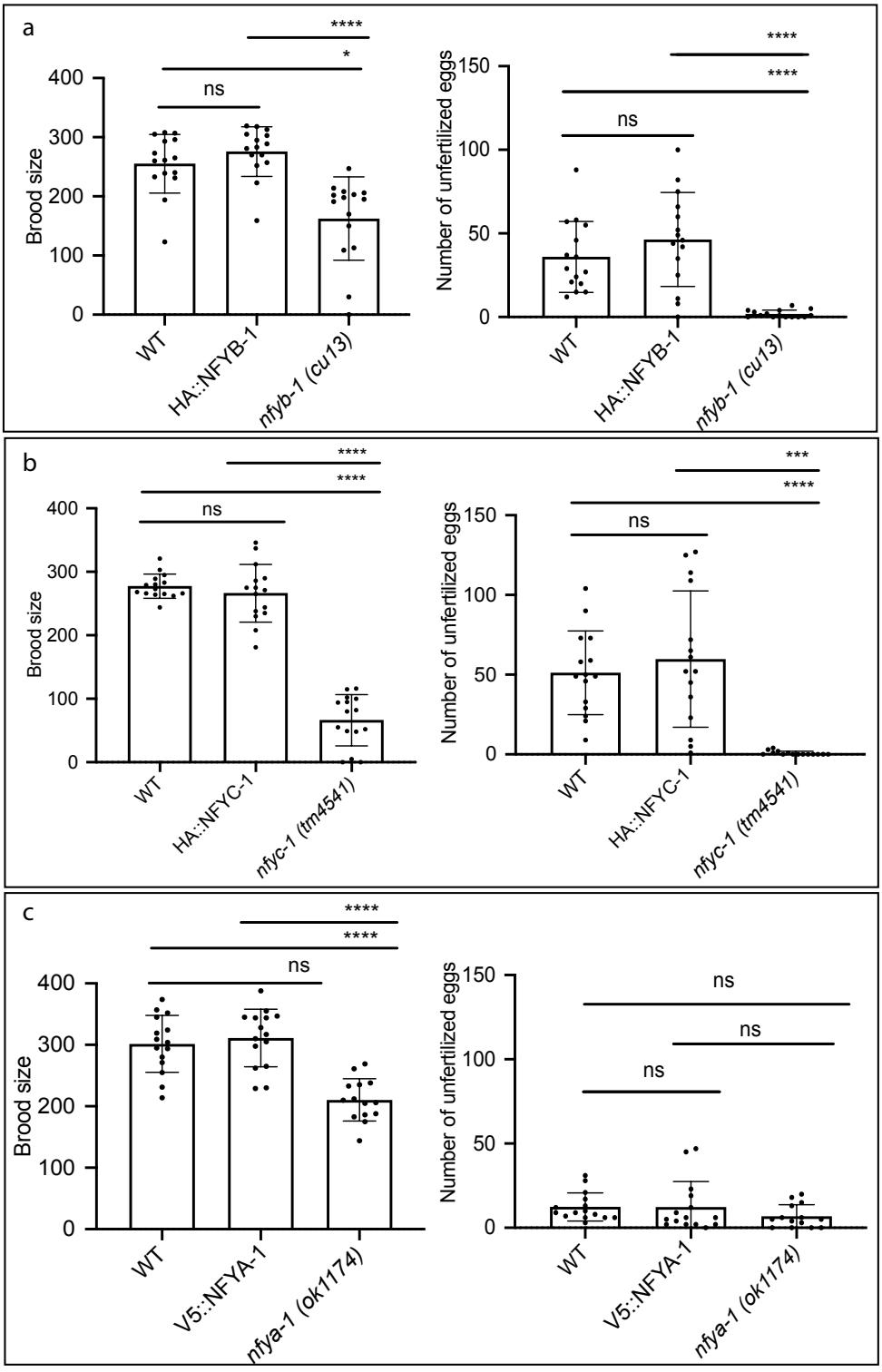

Figure 5.5. Functional assays of CRISPR-tagged NFY-A/B/C protein function. A, B and C - Brood size assay for the HA::NFYB-1, HA::NFYC-1 and V5::NFYA-1 (graph A, B and C respectively) compared to wild type and *nfyb-1(cu13), nfyc-1(tm4541) and nfya-1(ok1174)* mutant worms (respectively) over a period of 6 days. The left graphs show the number of progeny produced during this period (brood size). The right graphs show the number of unfertilized eggs. Data is expressed as mean ± SEM. Statistical significance was assessed using one-way ANOVA multiple comparison with Tukey's test. n=15, ns, not significant, *p<0.01, ***p<0.003, ****p<0,0001.

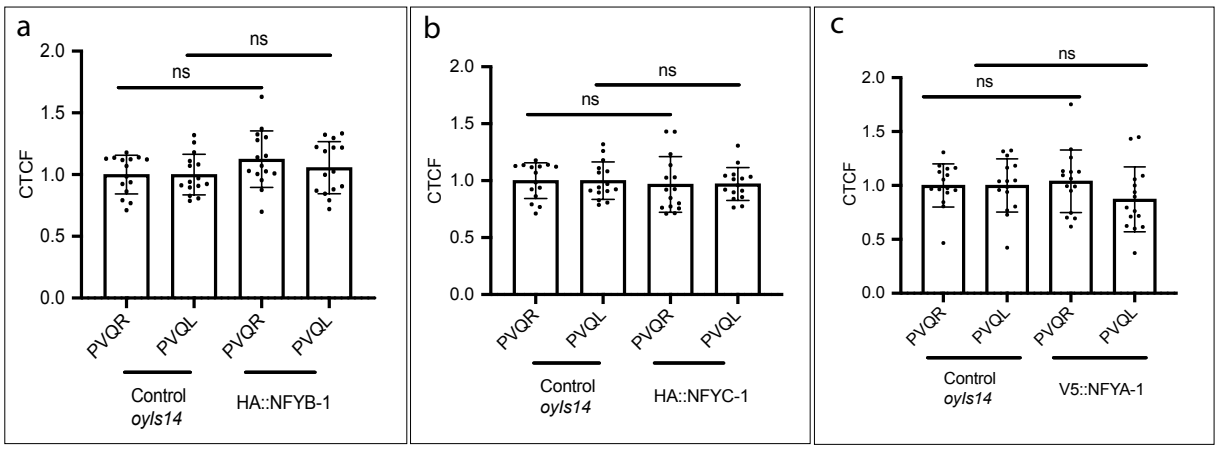

Figure 5.6. Assay to test HA::NFYB-1, HA::NFYC-1 and V5::NFYA-1 protein functionality. A – Quantification of expression (CTCF = GFP intensity) in the PVQ neurons for the control *oyIs14* (*Psra-6::GFP)* and HA::NFYB-1 strains. **B –** Quantification of expression (CTCF = GFP intensity) in the PVQ neurons for the control *oyIs14* (*Psra-6::GFP)* and HA::NFYC-1 strains. **C –** Quantification of expression (CTCF = GFP intensity) in both PVQs for the control *oyIs14* (*Psra-6::GFP)* and V5::NFYA-1 strain. Data is expressed as mean ± SD. Statistical significance was assessed using one-way ANOVA multiple comparison with Sidak's test. n=15 neurons; ns, not significant.

Taken together, these results demonstrate that epitope-tagging of the NFYA-1, NFYB-1 and NFYC-1 subunits does not affect global (brood size) and cell specific (PVQ specification) functions.

5.4.3. NFY complex subunit pulldown

To perform proteomics and ChIP-PCR/ChIP-seq analyses, the tagged proteins need to be immunoprecipitated and detected by western blot. To that end, I had to optimise the protein extraction method. I detected the HA::NFYB-1 subunit (46 kDa) band that is not detected in the wild type control (Figure 5.7a). However, I was not able to detect the two HA::NFYC-1 isoforms (isoform a – 26.7 kDa, isoform b – 22.7 kDa). I hypothesize that a higher protein concentration or optimization of the protein extraction may be needed. I also performed western blot analysis for the V5::NFYA-1, but the results were inconclusive (Figure 5.7b). In Figure 5.7, we observe three specific bands in the V5::NFYA-1 sample that are not present in the wild type control samples.

However, only one isoform (53.6 kDa) detected in this western has been described so far and there are two more bands present. These results could be due to post-translational modification of the NFYA-1 subunit leading to differential migration on the gel.

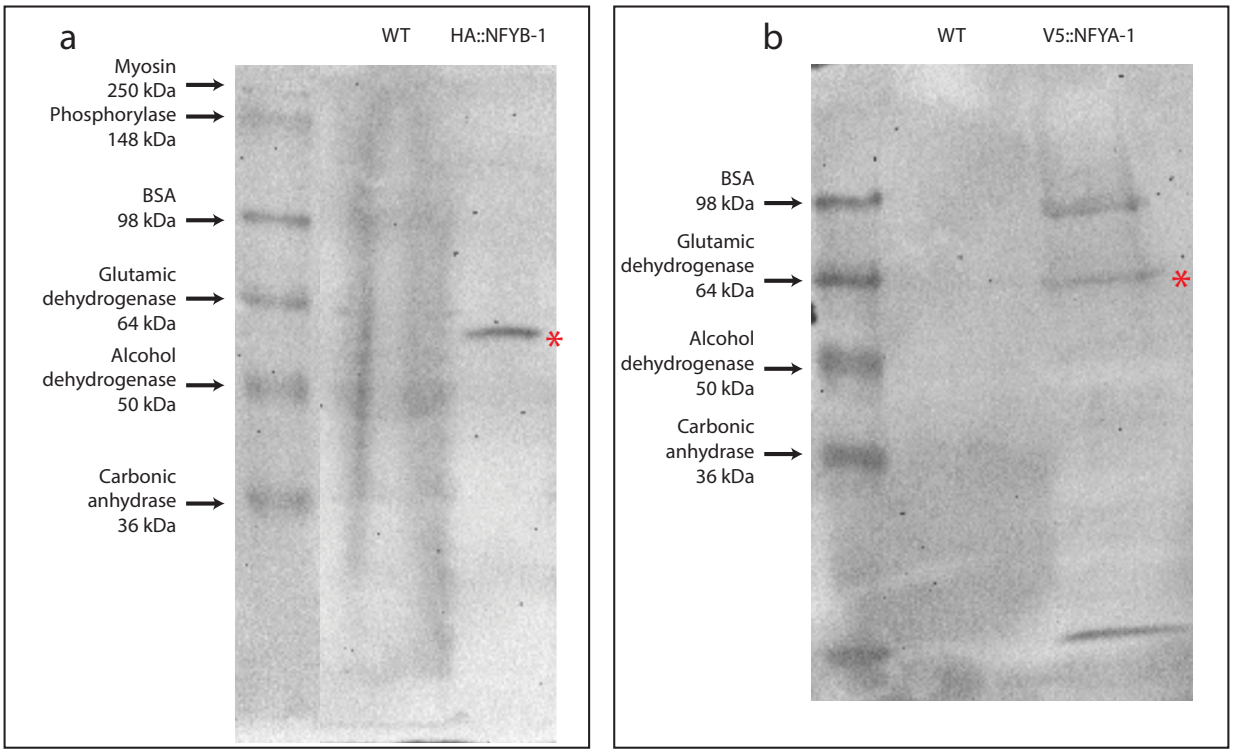

Figure 5.7. Western blot analyses showing HA::NFYB-1 and V5::NFYA-1 detection. NFY complex proteins were extracted from mixed staged worm population for V5::NFYA-1, HA::NFYB-1 and wild type control. Samples were run in a Western blot using anti-HA (A) and anti-V5 (B) antibodies, from Roche and Biorad respectively. **A –** Western blot to detect the HA::NFYB-1 protein (* expected size = 46 kDa). **B -** Western blot shows the HA::NFYA-1 protein detection (* expected size = 53.6 kDa).

After confirming that the HA::NFYB-1 protein could be detected by western blot, I carried out immunoprecipitation (IP) to detect the HA-tag in HA::NFYB-1 and wild type worms. The results clearly showed a band of 46 kDa in the HA::NFYB-1 sample that was not present in controls samples (wild type and HA::NFYB-1 with beads but no HA antibody), corresponding to the size of the NFYB-1 protein (Figure 5.8).

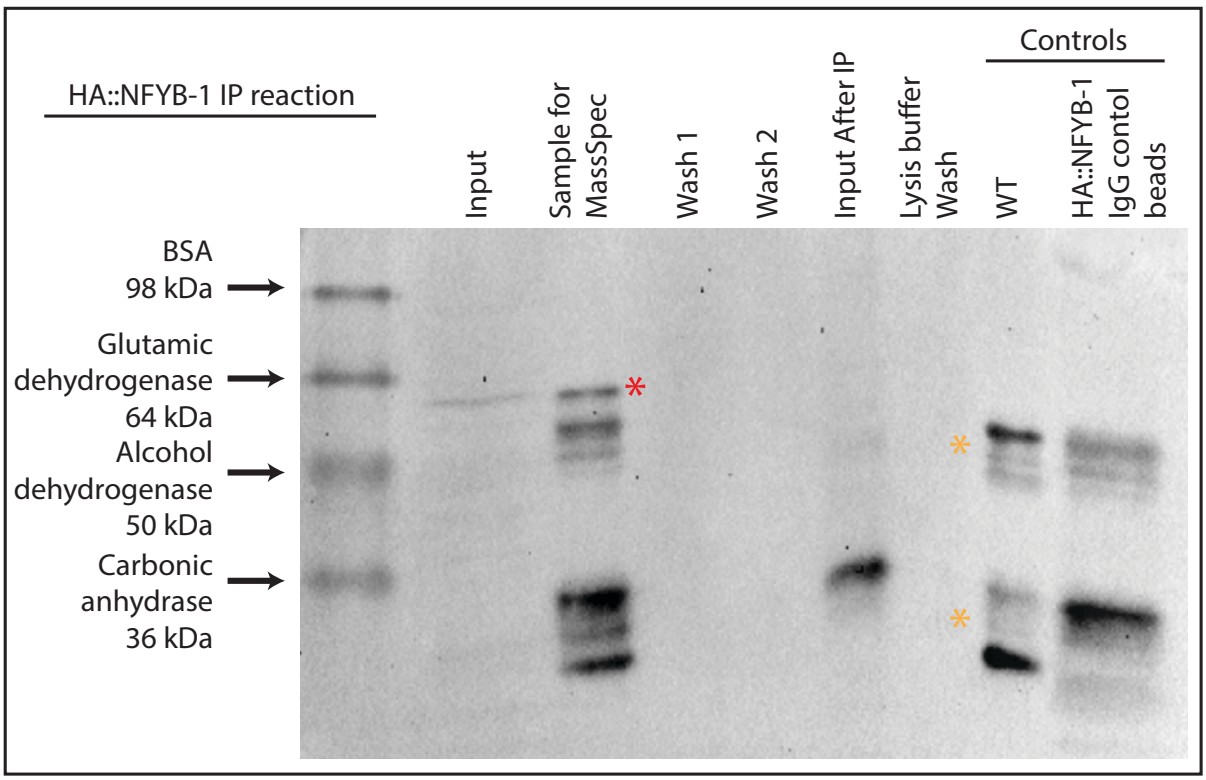

Figure 5.8. HA::NFYB-1 Immunoprecipitation reaction using magnetic beads. In this protocol, an HA antibody (Roche) and Dynobeads (Thermofisher) were used to isolate HA::NFYB-1 and possible interactors to send for MassSpec analysis. In the Western blot the HA::NFYB-1 band is only present in the Input and Sample for MassSpec lanes and not present in any of the control lanes. The Red (*) represents the NFYB-1 protein (46 kDa) and the orange (*) labels both IgG chains - heavy chain (50 kDa) and light chain (25 kDa).

Based on these promising results I sent one replicate sample for mass spectrometry analysis (data not shown). This pilot experiment detected both NFYB-1 and NFYC-1 confirming that the immunoprecipitation protocol worked and that other interactors detected might be good candidates for future work. However, the mass spectrometry analysis also detected unspecific proteins binding to the beads in the control samples. Perhaps, the wash steps were not stringent enough to avoid unspecific binding of proteins to the magnetic beads and further optimization of the protocol is required.

5.5. Discussion

C. elegans is regarded has one of the best models to study neuronal development due to its small and easy access to the nervous system. However, due to the low availability of antibodies that efficiently work in *C.* elegans, such as antibodies specific for the NFY complex, data resulting from proteomics (immunoprecipitation assay, western blot and mass-spectrometry) is often unsuccessful and/or limited. To overcome this issue and to better understand the NFY complex proteome, I endogenously tagged the NFY subunits (A, B, and C) using CRISPR technology. This was the first time that someone attempted to tag the NFYA/B and C subunits without disrupting protein function in a multicellular organism. The tagging of the NFY complex subunits will allow mass spectrometry analyses to be conducted, since GFP, 3xFLAG, HA, and V5 antibodies are available. Hence, this will allow me to identify novel co-activator/repressors that work together with the NFY complex to control gene expression.

In this study, I first attempted to tag the NFY subunits with GFP since it would allow me to follow protein localization with ease and perform immunoprecipitation at the same time. This was however unsuccessful since larger tags present lower CRISPR-Cas9 insertion efficiency as mentioned in previous studies (Mikuni *et al.*, 2016). Consequently, I tried to insert a smaller tag in size (3xFLAG-tag). However, this approach also proved to be unsuccessful possibly due to the size of the tag - GFP (870 bp) and 3xFLAG (75 bp). Previous studies performed in yeast demonstrated that different sizes of tags can affect protein function at different levels (Andresen *et al.*, 2004). Andresen *et al.* (2004) tagged the *TUB2* gene with different size tags, 1x TetCys motif (10 aa), 2x TetCys motif (20 aa), 3x TetCys-motif (29 aa), or GFP (244 aa), and demonstrated that even small changes in tag size can disrupt protein function. Further, the authors were able to easily tag the *TUB2* gene with 1x TetCys motif (10 aa/30 bp) and 2x TetCys motif (20 aa/60 bp), while maintaining protein functionality. However, they were not able to tag the *TUB2* gene with 3x TetCys-motif (29 aa/87 bp) or GFP (244 aa/735 bp), and maintain protein functionality (Andresen *et al.*, 2004).

Moreover, other studies have demonstrated that larger tags, such as all fluorescent proteins, can easily disrupt protein structure, function and location within the cell (Koch *et al.*, 2018; Moore and Murphy, 2009). Altogether, these observations indicate that small changes in tag size are a crucial element to take into account when performing gene editing for protein tagging. The limitation presented in this study reinforced the idea that it is important to perform proper functional validation after protein tagging, since incorrect protein tagging can result in impaired protein function or loss of function. A possible approach to functionally validate the tagged proteins is by comparing the tagged protein transgenic strain with a previously described loss-of-function mutant strain and a wild-type strain (Andresen *et al.*, 2004; Moore and Murphy, 2009). To accomplish this validation step, I compare brood size and PVQ GFP expression (*oyIs14* reporter) in wild type, tagged proteins (V5::NFYA-1, HA::NFYB-1 and HA::NFYC-1), and loss-of-function mutant *(nfya-1(ok1174), nfyb-1(cu13) and nfyc-1(tm4195)*. All the CRISPR-tagged lines generated in this study (V5::NFYA-1, HA::NFYB-1 and HA::NFYC-1) presented brood-size and GFP expression levels in PVQ neurons comparable to wild type levels, indicating that the tagged proteins were functional. This was the first time that the NFY subunits were endogenously tagged in a multicellular organism.

Based on the results obtained in this study, I demonstrate that larger tags present a lower insertion efficiency likely due to their size. Therefore, smaller tags, such as HA (27 bp) or V5 (42 bp), are considered advantageous as they have higher insertion efficiency, less probability to affect gene expression and protein structure, and protein-protein or protein–DNA interaction (Zhang *et al.*, 2015; Zhao *et al.*, 2013). Moreover, I showed that introducing a single epitope tag (HA/V5) with CRISPR technology is sufficient to detect the NFY subunits by western blot and IP, avoiding the need to use specific antibodies against the proteins of interest (NFY subunits). So far, preliminary mass-spectrometry results for the HA::NFYB-1 subunit are promising and detect both the NFYB-1 and the NFYC-1 subunits. In the future, additional samples will be collected and optimization of the IP protocol performed to enable mass spectrometry and ChIP-PCR/ChIP Sequencing experiments.

Chapter 6. Final Discussion

Achieving correct brain architecture is an exceptionally complex process requiring precisely controlled cell specification, migration, axon outgrowth and axon guidance. It is well known that for all these processes to occur specific genes need to be expressed/repressed in precise spatial and temporal patterns to drive molecular programs for each neuron. My work aimed to identify new molecular mechanisms that drive neuronal development. Specifically, I focussed on the PVQ neurons since very little is known about the formation and specification of these neurons.

6.1. Characterization of neuron-specific transcriptional elements on the promoter of the PVQ-expressed *sra-6* gene in C. elegans

Here, I performed genetic and cis-regulatory analysis to identify motifs within the *sra-6* promoter that are responsible for gene expression in specific neurons. By dissecting the *sra*-6 promoter, I identified the promoter region (OPM2.5 fragment) that is required to control *sra-6* expression specifically in the PVQs. However, transgenic worms expressing the OPM2.5 region presented PVQ axonal defects. Although the origin of the PVQs defects is not clear, I hypothesize that an interaction between the OPM2.5 fragment and the *let-858* 3' UTR (also present in the reporter plasmid) may be responsible for these defects (Figure 6.1a). This hypothesis is based on the fact that reducing the size of the OPM2.5 fragment, by removing different binding sites, did not circumvent the PVQ defects (Figure 6.1a). Moreover, injection of the empty pPD117.01 (no promoter + GFP + *let-858* 3' UTR) and of the full *Psra-6::GFP* (pPD117.01 with *let-858* 3' UTR) constructs did not cause any detectable defects (Figure 6.1c & d). Thus, the presence of pPD117.01 plasmid alone or with the full *sra-6* promoter is not sufficient to cause the PVQ defects (Figure 6.1c & d). In addition, when I tried to inject the OPM2.5 fragment alone into the *oyIs14* transgenic strain or to inject in a different plasmid (with OPM2.5::*GFP*::*unc-54* 3' UTR) into wild-type

worms, and try to remove the PVQ defects, I was not able to detect any axonal defects or GFP expression, respectively (Figure 6.1e).

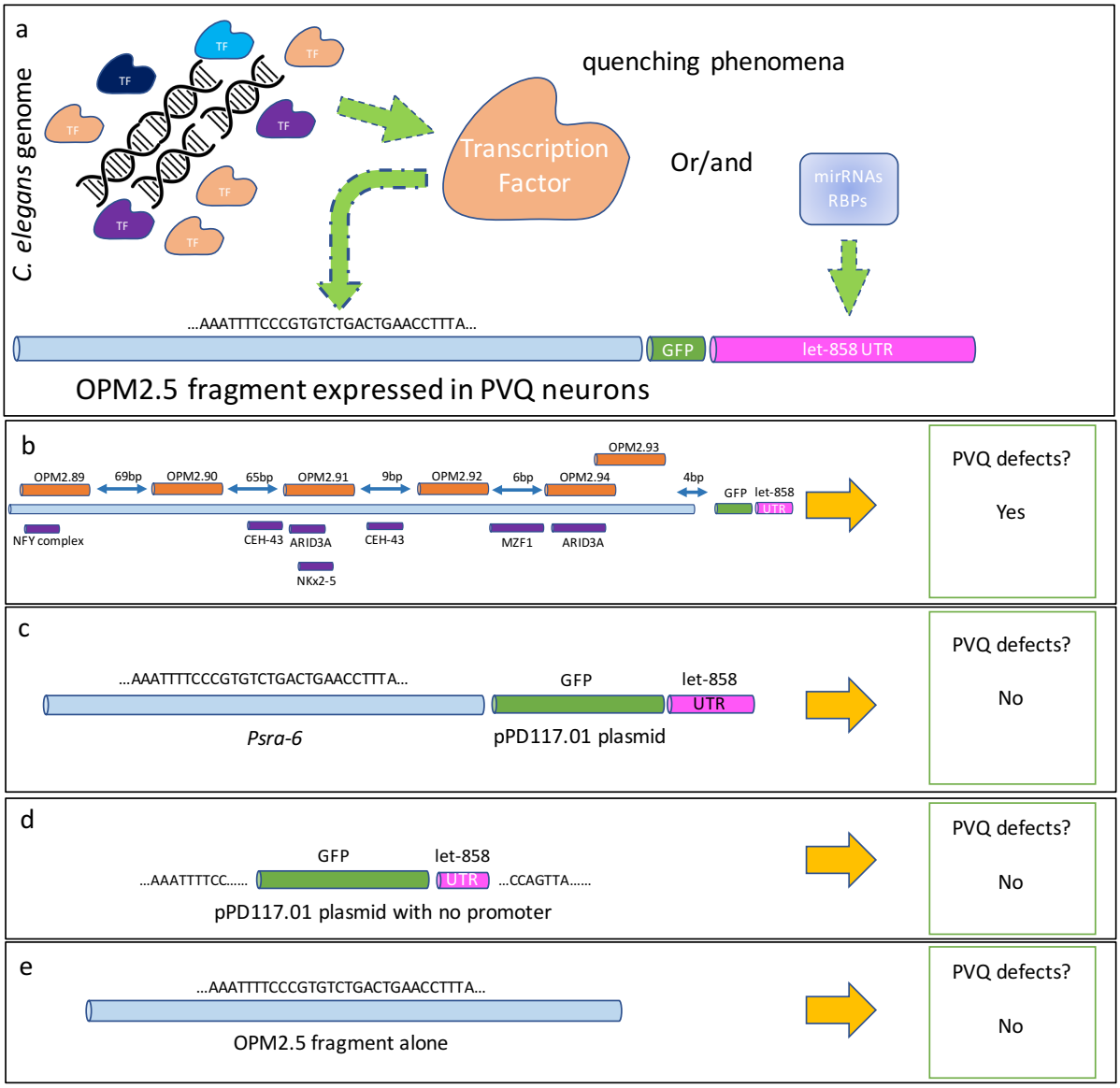

Figure 6.1. Schematic representation of the quenching phenomena affecting the promoter characterization experiments and the location of the putative binding sites for the different TF. **A-** The images represent a TF being quenched by being directed to the OPM2.5 fragment and miRNAs/RNA binding proteins (RBP) binding to the *let-858* 3'UTR leading to the PVQ defects. **B –** This scheme represents the binding sites present in a smaller fragment (OPM2.89F) that presents 100% of PVQ defects. The most prominent candidates that could cause these defects are the NFY complex, CEH-43, Arid3a (CFI-1 in *C. elegans*), NK2-5 (CEH-22 in *C. elegans*) and MZF1. The orange cylindrical shapes - OPM2.89 to OPM2.93, represent the binding site for the primers used. **C –** Injection of full *sra-6* promoter (driving the GFP together with the *let-858* 3' UTR) do not cause PVQ defects. **D –** Injection of the pPD117.01 does not cause any PVQ defects. **E –** Injection of the OPM2.5 fragment it is not responsible for PVQ defects.

Taken together, these findings suggest that the OPM2.5 fragment needs to be associated with the *let-858* 3' UTR to express GFP and to disrupt PVQ development. Previous studies in *C. elegans* demonstrated that injection of specific constructs or co-injection markers can lead to formation of artefacts like artificial fluorescence in some cells, such as in muscle cells and in the PVT neurons (Boulin *et al.*, 2006). These observations suggested that the 3' UTR alone could be the source of artefact formation, including unspecific GFP expression (Boulin *et al.*, 2006). Another possible explanation, is that interaction between different sequences, such as interaction between promoters, reporter genes and 3'UTRs, can result in unspecific findings. For example, a study performed by Jin *et al.* (2019) using *C. elegans* revealed that injection of the *rol-6* vector in conjunction with an empty GFP construct (absent promoter region) leads to unspecific GFP expression in the CP09 neurons. This study suggested that specific UTR sequences combined with vectors or other sequences can cause background GFP expression, and particularly in this study due to homologous recombination of *rol-6* co-injection plasmid and the GFP construct (Jin *et al.*, 2019). Based on these observations, it is possible that injection of the OPM2.5 full-length fragment and/or shorter fragments in conjunction with the *let-858* 3'UTR may cause the PVQ specific GFP expression and the unwanted PVQ defects observed in this experiment (Figure 6.1a). However, it is not clear whether these defects arise from binding of specific TFs to the OPM2.5, binding of microRNAs to the 3' UTR, or both, ultimately leading to a quenching event that disrupts PVQ development. A possible future approach is to inject the OPM2.5 fragment + GFP + *unc-54 3' UTR* into *oyIs14* transgenic worms and examine whether there are PVQ defects. If no defects are detected it demonstrates that the OPM2.5 needs to work together with the *let-858 3' UTR* to originate PVQ defects. Another possible experiment would be to screen for miRNA binding sites in the *let-858* 3' UTR and mutate those sites to examine if the defects disappear.

6.2. The role of the NFY transcription factor complex in neuronal development

In chapter 3, dissection of the OPM2.5 promoter revealed transcription factor binding motifs that may be relevant for the control of *sra-6* expression in the PVQs, such as CFI-1, CEH-43, and NFY transcription factors (Figure 6.1b). Some of these TFs are known to be required for neuron formation in *C. elegans* and mouse models (Aspock and Burglin, 2001; Doitsidou *et al.*, 2013; Shaham and Bargmann, 2002; Yamanaka *et al.*, 2016). By analysing mutant worms for these TFs, I was able to identify the NFY complex as a strong candidate to be involved in PVQ development. Based on the analysis performed on chapter 4, I showed that the NFY complex cell autonomously controls PVQ gene expression, possibly by binding to CCAAT conserved motifs present in the promoter regions, and control PVQ specification and axon guidance but not PVQ generation or maintenance (Figure 6.2).

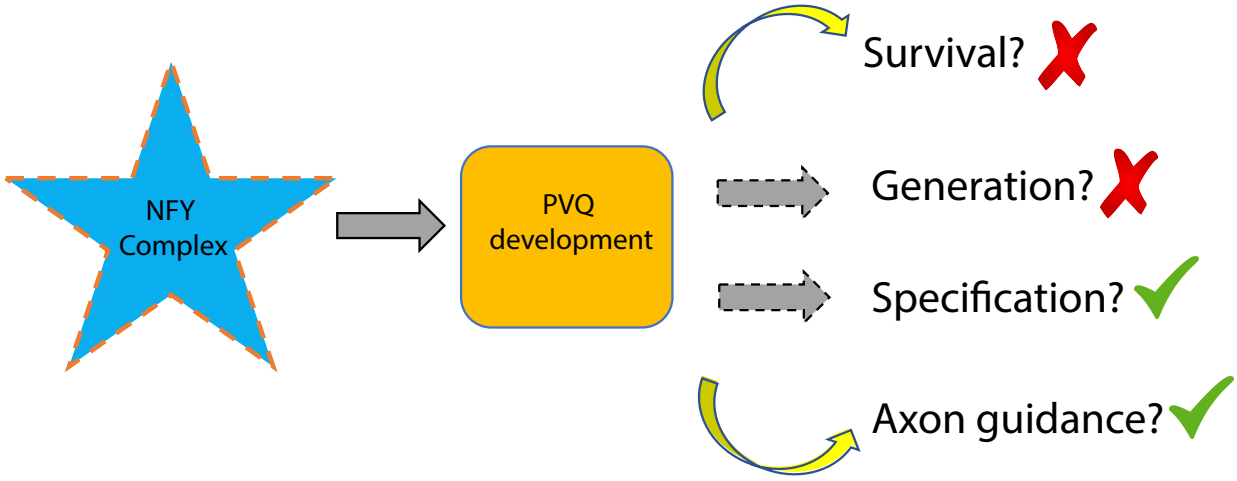

Figure 6.2. Schematic representation of the NFY complex function during PVQ development. The NFY complex is essential for PVQ differentiation and axon guidance but not important for PVQ generation or survival.

6.2.1 Dissecting the NFY complex role in transcriptional regulation using *C. elegans*

The role of each subunit of the NFY complex remains poorly understood. In this study, I showed that double mutant worms for NFYB/NFYC subunits present the same percentage of PVQ defects when compared to single NFYB or NFYC worms. Interestingly, NFYA-1 mutant worms showed a more severe neuronal phenotype when compared with single and double NFYB-1 and NFYC-1 mutant worms. These findings suggest that the NFYA-1 subunit may have an independent role from the NFYB-1 and NFYC-1 subunits (Figure 6.3). This is the first evidence of a potential independent role of the NFYA-1 subunit, and suggests that other proteins may be able to perform similar functions to the NFYB-1 and NFYC-1 subunits, and/or that the NFYA-1 subunit alone is sufficient to perform basal level regulation during neuronal development (Figure 6.3a). In addition, double mutants for *nfya-1* and *nfya-2* exhibited a higher percentage of PVQ neurons without GFP when compared with the *nfya-1* single mutant. These results support the idea that the different NFYA subunits act independently in order to regulate gene expression in *C. elegans* (Figure 6.3b & c). Previous studies performed by Deng *et al.* (2007) showed that the NFYB/NFYC dimer is able to interact with both NFYA-1 and NFYA-2 subunits to form two independent complexes. However, it is not clear whether these two complexes compete for the binding of the same CCAAT motif within the promoter regions or have independent binding sites. A possible future experiment is to perform Yeast one-hybrid for both NFY complexes (NFYB/NFYC/NFYA-1 and NFYB/NFYC/NFYA-2) and assess whether they are able to bind to the same sequences. Another possible approach is to perform ChIP-PCR by pulling down the NFYA-1 or the NFYA-2 subunits, analysing if there is an overlap between binding regions.

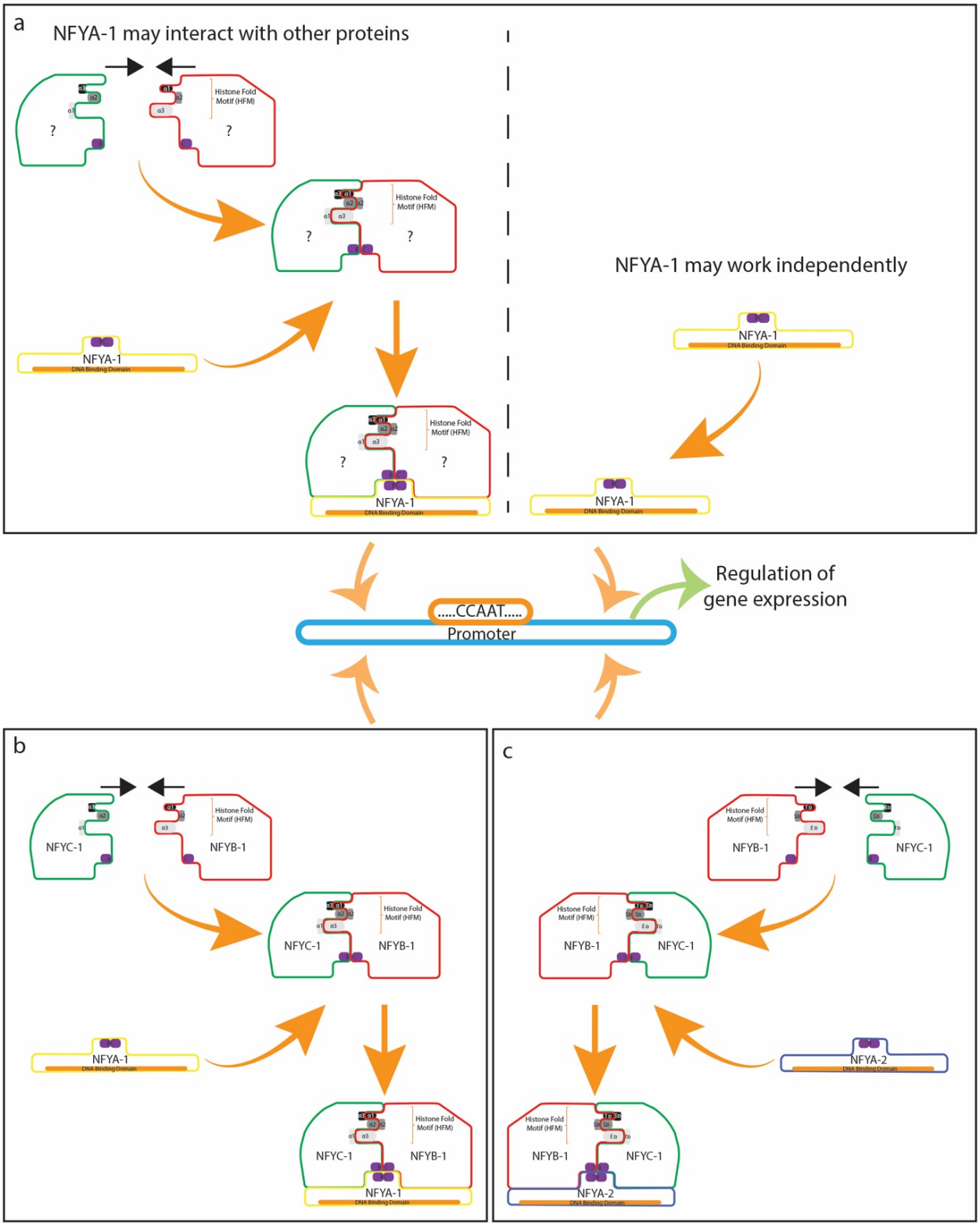

Figure 6.3. NFY complex possible action mechanism. A – NFYA-1 may interact with other proteins to regulate gene expression (left model) or NFYA-1 may work alone to regulate transcription (right model). **B –** Established NFYB-1/NFYC-1/NFYA-1 trimer complex formation. **C –** NFYB-1/NFYC-1/NFYA-2 trimer complex formation.

In this study, I hypothesized that the NFY complex could be required for different steps of PVQ development: generation, differentiation and/or survival. To test this hypothesis, I analysed several reporters expressed in the PVQ neurons. The results showed that the NFY complex plays a pivotal role in directly or indirectly regulating most PVQ expressed genes, and therefore, in promoting PVQ neuronal fate (Figure 6.4). To elucidate whether the NFY complex directly regulates PVQ gene expression, I mutated the CCAAT motif of NFYA-1-regulated genes to abrogate NYFA-1 binding. I found that intact CCAAT sites are required for gene expression in the PVQ neurons, suggesting that the NFY complex directly binds to these sites. This analysis confirms previous studies performed by Deng *et al.* (2007) that revealed that any change in these CCAAT motifs will impair the ability of the NFY complex to bind to the conserved motifs. To confirm that the NFY complex directly regulates the expression of these genes, ChIP-PCR analyses will be performed in future experiments.

Previous studies have shown that the NFY complex is required for maintenance of different neuronal subsets in mammals, including striatal medium spiny neurons (MSNs), cerebellar Purkinje cells, and spinal/brainstem motor neurons (Yamanaka *et al.*, 2016). Consequently, the NFY complex may have a broad role in the development of the nervous system. Bioinformatic analysis performed on ChIP-Seq data from endogenously-tagged NFYA-1 in *C. elegans* (modencode.org) identified putative binding sites in the promoter region of pan-neuronal genes: *rab-3, unc-11,* and *ric-4.* These data suggest that the NFY complex regulates general neuronal features in multiple neuron subtypes. To test this hypothesis, I analysed the function of the NFY complex in a different neuronal subtype (motor neurons) by examining the expression of pan-neuronal markers (*Punc-11*, *Prab-3, ric-4* fosmid) (Figure 6.4). My data revealed that *nfya-1* is required for correct expression of promoter-based reporters for *unc-11* and *rab-3* but not for a *ric-4* fosmid reporter (Figure 6.4). One possible explanation for these results is that different TFs are required for the regulation of pan-neuronal genes in the different neurons through intricate transcriptional networks. This potential regulatory model was demonstrated by Stefanakis *et al.* (2015), where they

identified non-overlapping elements present in the *ric-4* and *snb-1* pan-neuronal reporters that create overlapping expression patterns independently from each other. Therefore, different regulatory elements may have redundant functions and act in parallel with NFY complex elements (CCAAT boxes) to control the *ric-4* expression. To further explore this possibility, *unc-11* and *rab-3* fosmid reporters could be analysed to test for the presence of additional regulatory elements outside the promoters of *unc-11* and *rab-3*. Moreover, as these genes encode synaptic proteins, neurotransmission could be affected and therefore animal behaviour may be impaired. To investigate whether neuronal function is in fact compromised, motility assays could be performed in future experiments to analyse *C. elegans* motor function. Another possible approach to examine neuronal function would be to use a calcium sensor reporter expressed in the motor neurons to directly measure neuronal activity in *nfya-1* mutant animals.

In summary, by mutating the NFY complex or the ability of the NFY complex to interact with the promoter of target genes, we impaired the expression of neuron-specific and pan-neuronal features of *C. elegans* interneurons (PVQs) and motor neurons (Figure 6.4).

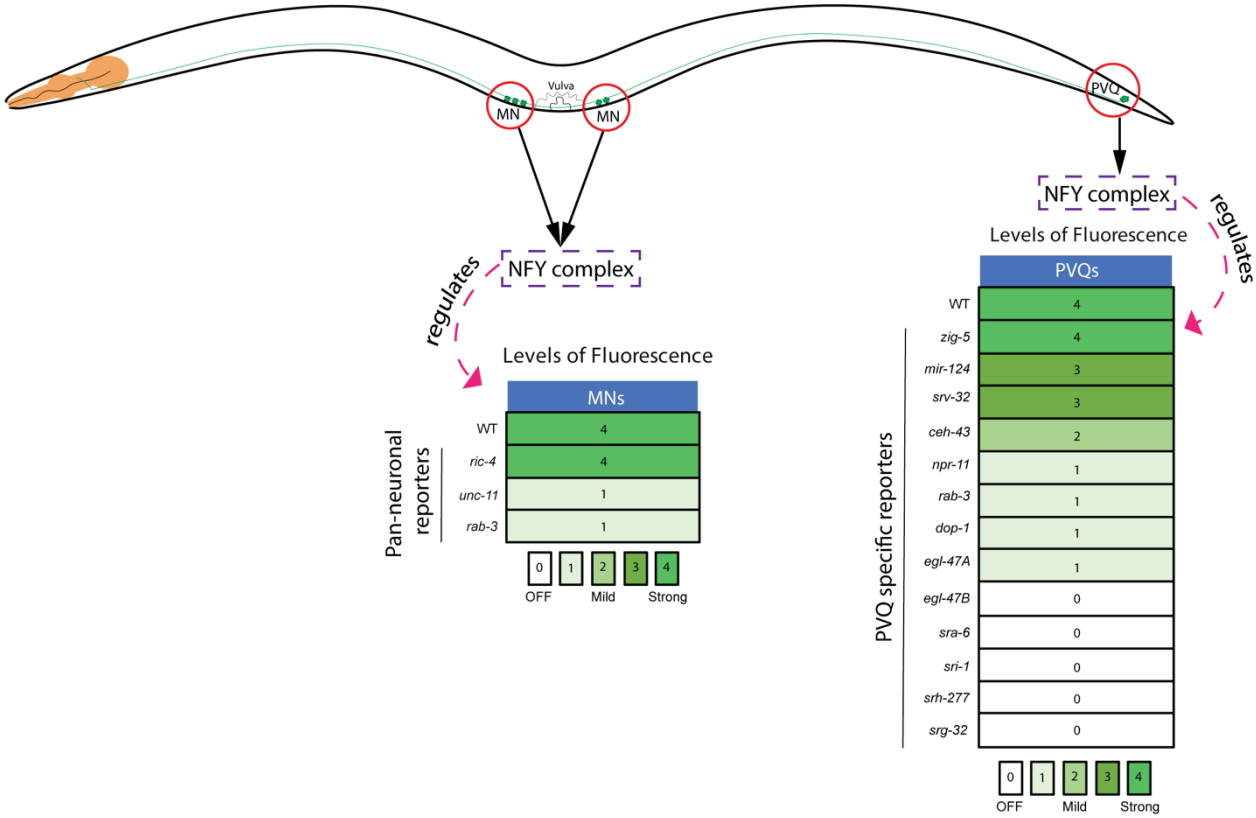

Figure 6.4. Schematic representation of transcriptional regulation levels during neurogenesis for the different genes analysed in this study. Dynamic TF networks responsible for acquisition of neuronal identity, such as expression of GPCRs, neuropeptides, and their receptors (right). Disruption of the NFY complex leads to the loss of specific PVQ neuron fate reporters (right). The NFY complex regulates expression of pan-neuronal genes required for synaptic transmission in motor neurons (left).

6.2.2. The NFY complex interactome in neuronal development

Previous studies revealed that the NFY complex needs to interact with activators or repressors of transcription, such as p300 and ATF6, to perform its function (Li *et al.*, 1998; Luo *et al.*, 2008). Moreover, based on my analysis and results it appears that the NFY complex works in concert with specific activators or repressors to regulate the same genes differently in specific neuronal subtypes. For example, in chapter 5, I demonstrate that the NFY complex regulates the *sra-6* gene in the PVQs but not in

the ASH neurons. This difference in regulation could be due to specific interaction of the NFY complex with a PVQ specific NFY complex co-activator.

This information led me to hypothesize that the NFY complex works in concert with other proteins to regulate different genes required for the development of specific neuronal subsets (Figure 6.5). This model supports the idea that, despite the NFY complex being broadly expressed across the *C. elegans* nervous system, the NFY complex will behave in a specific manner in different cells due to interaction with proteins that can be only expressed in specific neuronal subsets.

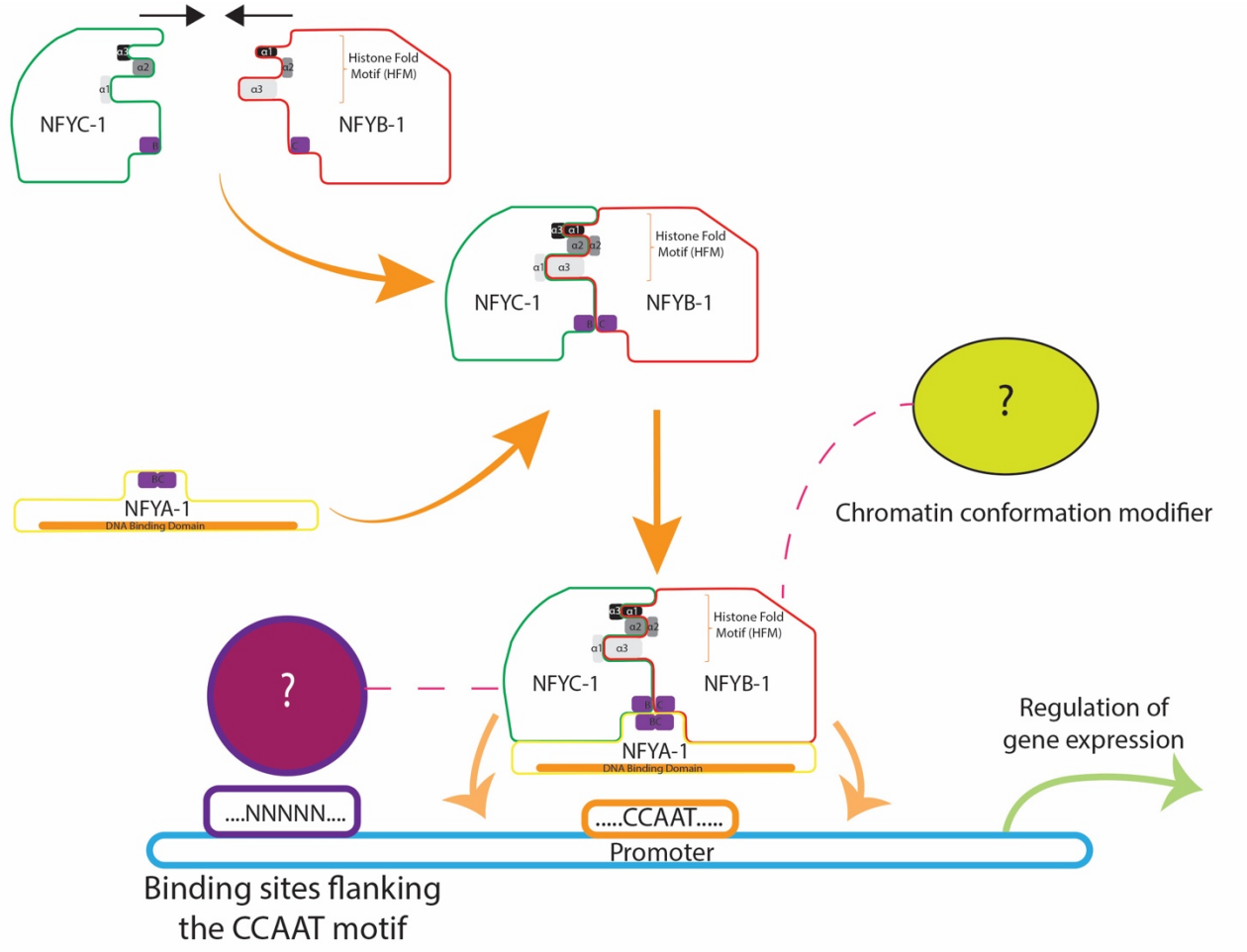

Figure 6.5. Representation of NFY complex and co-activators/repressors interaction. This image shows possible interactions between the NFY complex and other coactivators/corepressors of transcription. This model aims to explain why specific genes can be regulated by the NFY complex in specific neuronal types and not in others.

To explore this model and identify which proteins could be relevant for the NFY complex neuronal function I endogenously tagged HA::NFYB-1, HA::NFYC-1 and V5::NFYA-1 subunits. All these newly genetically engineered NFY subunits were functionally validated and two of them were detected by western blot. In the future, immunoprecipitation and mass spectrometry analyses will be performed with these CRISPR lines to identify putative co-activators and repressors of transcription. This will provide a more in-depth view on how the NFY complex is able to regulate different genes in different neuronal types.

6.4. Concluding remarks

Overall, I show that the NFY complex is required to control the expression of PVQs specific genes that are required for PVQ identity (Figure 6.5). Additionally, I show that the NFY complex regulates pan-neuronal gene expression, suggesting that the NFY complex has a wider neuronal function. Further analyses are warranted to understand how is the NFY complex able to perform different functions in different cell subtypes. To explore the model that the NFY complex interacts with different co-activator in different cells types, leading to different levels of transcription regulation, proteomics analysis will be performed to identify possible co-activators or repressors involved in neuronal development in specific neuronal subsets.

7. References

Altun, Z.F., Herndon, L.A., Wolkow, C.A., Crocker, C., Lints, R. , and Hall, D.H. (2002-2020). WormAtlas.

Andresen, M., Schmitz-Salue, R., and Jakobs, S. (2004). Short tetracysteine tags to beta-tubulin demonstrate the significance of small labels for live cell imaging. Mol Biol Cell *15*, 5616-5622.

Anitha, A., Nakamura, K., Yamada, K., Suda, S., Thanseem, I., Tsujii, M., Iwayama, Y., Hattori, E., Toyota, T., Miyachi, T.*, et al.* (2008). Genetic analyses of roundabout (ROBO) axon guidance receptors in autism. Am J Med Genet B Neuropsychiatr Genet *147B*, 1019-1027.

Arents, G., and Moudrianakis, E.N. (1995). The histone fold: a ubiquitous architectural motif utilized in DNA compaction and protein dimerization. Proc Natl Acad Sci U S A *92*, 11170-11174.

Aspock, G., and Burglin, T.R. (2001). The Caenorhabditis elegans distal-less ortholog ceh-43 is required for development of the anterior hypodermis. Dev Dyn *222*, 403-409.

Aurelio, O., Boulin, T., and Hobert, O. (2003). Identification of spatial and temporal cues that regulate postembryonic expression of axon maintenance factors in the C. elegans ventral nerve cord. Development *130*, 599-610.

Aurelio, O., Hall, D.H., and Hobert, O. (2002). Immunoglobulin-domain proteins required for maintenance of ventral nerve cord organization. Science *295*, 686-690.

Baumeister, R., Liu, Y., and Ruvkun, G. (1996). Lineage-specific regulators couple cell lineage asymmetry to the transcription of the Caenorhabditis elegans POU gene unc-86 during neurogenesis. Genes Dev *10*, 1395-1410.

Baxevanis, A.D., Arents, G., Moudrianakis, E.N., and Landsman, D. (1995). A variety of DNA-binding and multimeric proteins contain the histone fold motif. Nucleic Acids Res *23*, 2685-2691.

Benatti, P., Chiaramonte, M.L., Lorenzo, M., Hartley, J.A., Hochhauser, D., Gnesutta, N., Mantovani, R., Imbriano, C., and Dolfini, D. (2016). NF-Y activates genes of metabolic pathways altered in cancer cells. Oncotarget *7*, 1633-1650.

Boulin, T., Etchberger, J.F., and Hobert, O. (2006). Reporter gene fusions. WormBook, 1-23.

Braitenberg, V., and Atwood, R.P. (1958). Morphological observations on the cerebellar cortex. The Journal of Comparative Neurology *109*, 1-33.

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

Brooks, M.D., Cirrone, J., Pasquino, A.V., Alvarez, J.M., Swift, J., Mittal, S., Juang, C.L., Varala, K., Gutierrez, R.A., Krouk, G.*, et al.* (2019). Network Walking charts transcriptional dynamics of nitrogen signaling by integrating validated and predicted genome-wide interactions. Nat Commun *10*, 1569.

Bülow, H.E., Boulin, T., and Hobert, O. (2004). Differential Functions of the C. elegans FGF Receptor in Axon Outgrowth and Maintenance of Axon Position. Neuron *42*, 367-374.

Cameron, S., Clark, S.G., McDermott, J.B., Aamodt, E., and Horvitz, H.R. (2002). PAG-3, a Zn-finger transcription factor, determines neuroblast fate in C. elegans. Development *129*, 1763-1774.

Charron, F., Stein, E., Jeong, J., McMahon, A.P., and Tessier-Lavigne, M. (2003). The Morphogen Sonic Hedgehog Is an Axonal Chemoattractant that Collaborates with Netrin-1 in Midline Axon Guidance. Cell *113*, 11-23.

Clark, A.M., Goldstein, L.D., Tevlin, M., Tavare, S., Shaham, S., and Miska, E.A. (2010). The microRNA miR-124 controls gene expression in the sensory nervous system of Caenorhabditis elegans. Nucleic Acids Res *38*, 3780-3793.

Clark, S.G., and Chiu, C. (2003). C. elegans ZAG-1, a Zn-finger-homeodomain protein, regulates axonal development and neuronal differentiation. Development *130*, 3781-3794.

Cook, S.J., Jarrell, T.A., Brittin, C.A., Wang, Y., Bloniarz, A.E., Yakovlev, M.A., Nguyen, K.C.Q., Tang, L.T., Bayer, E.A., Duerr, J.S.*, et al.* (2019). Whole-animal connectomes of both Caenorhabditis elegans sexes. Nature *571*, 63-71.

Deng, H., Sun, Y., Zhang, Y., Luo, X., Hou, W., Yan, L., Chen, Y., Tian, E., Han, J., and Zhang, H. (2007). Transcription factor NFY globally represses the expression of the C. elegans Hox gene Abdominal-B homolog egl-5. Dev Biol *308*, 583-592.

Dent, E.W., Gupton, S.L., and Gertler, F.B. (2011). The growth cone cytoskeleton in axon outgrowth and guidance. Cold Spring Harb Perspect Biol *3*.

Desai, C., Garriga, G., McIntire, S.L., and Horvitz, H.R. (1988). A genetic pathway for the development of the Caenorhabditis elegans HSN motor neurons. Nature *336*, 638-646.

Doitsidou, M., Flames, N., Topalidou, I., Abe, N., Felton, T., Remesal, L., Popovitchenko, T., Mann, R., Chalfie, M., and Hobert, O. (2013). A combinatorial regulatory signature controls terminal differentiation of the dopaminergic nervous system in C. elegans. Genes Dev *27*, 1391-1405.

Dokshin, G.A., Ghanta, K.S., Piscopo, K.M., and Mello, C.C. (2018). Robust Genome Editing with Short Single-Stranded and Long, Partially Single-Stranded DNA Donors in Caenorhabditis elegans. Genetics *210*, 781-787.

Durbin, R.M. (1987). Studies on the development and organisation of the nervous system of C. elegans.

Flames, N., and Hobert, O. (2009). Gene regulatory logic of dopamine neuron differentiation. Nature *458*, 885-889.

Frank, C.A., Baum, P.D., and Garriga, G. (2003). HLH-14 is a C. elegans achaete-scute protein that promotes neurogenesis through asymmetric cell division. Development *130*, 6507-6518.

Fried, M., and Crothers, D.M. (1981). Equilibria and kinetics of lac repressor-operator interactions by polyacrylamide gel electrophoresis. Nucleic Acids Res *9*, 6505-6525.

Gallo, G., and Letourneau, P.C. (2004). Regulation of growth cone actin filaments by guidance cues. J Neurobiol *58*, 92-102.

Garel, S., Garcia-Dominguez, M., and Charnay, P. (2000). Control of the migratory pathway of facial branchiomotor neurones. Development *127*, 5297-5307.

Gertz, J., Reddy, T.E., Varley, K.E., Garabedian, M.J., and Myers, R.M. (2012). Genistein and bisphenol A exposure cause estrogen receptor 1 to bind thousands of sites in a cell type-specific manner. Genome Res *22*, 2153-2162.

Gramstrup Petersen, J., Rojo Romanos, T., Juozaityte, V., Redo Riveiro, A., Hums, I., Traunmuller, L., Zimmer, M., and Pocock, R. (2013). EGL-13/SoxD specifies distinct O2 and CO2 sensory neuron fates in Caenorhabditis elegans. PLoS Genet *9*, e1003511.

Hallam, S., Singer, E., Waring, D., and Jin, Y. (2000). The C. elegans NeuroD homolog cnd-1 functions in multiple aspects of motor neuron fate specification. Development *127*, 4239-4252.

Hannula-Jouppi, K., Kaminen-Ahola, N., Taipale, M., Eklund, R., Nopola-Hemmi, J., Kaariainen, H., and Kere, J. (2005). The axon guidance receptor gene ROBO1 is a candidate gene for developmental dyslexia. PLoS Genet *1*, e50.

Hedgecock, E.M., Culotti, J.G., and Hall, D.H. (1990). The unc-5, unc-6, and unc-40 genes guide circumferential migrations of pioneer axons and mesodermal cells on the epidermis in C. elegans. Neuron *4*, 61-85.

Herrero, J., Muffato, M., Beal, K., Fitzgerald, S., Gordon, L., Pignatelli, M., Vilella, A.J., Searle, S.M., Amode, R., Brent, S.*, et al.* (2016). Ensembl comparative genomics resources. Database (Oxford) *2016*.

Hobert, O. (2011). Regulation of terminal differentiation programs in the nervous system. Annu Rev Cell Dev Biol *27*, 681-696.

Hu, Q., and Maity, S.N. (2000). Stable expression of a dominant negative mutant of CCAAT binding factor/NF-Y in mouse fibroblast cells resulting in retardation of cell growth and inhibition of transcription of various cellular genes. J Biol Chem *275*, 4435-4444.

Hume, M.A., Barrera, L.A., Gisselbrecht, S.S., and Bulyk, M.L. (2015). UniPROBE, update 2015: new tools and content for the online database of protein-binding microarray data on protein-DNA interactions. Nucleic Acids Res *43*, D117-122.

Hutter, H. (2003). Extracellular cues and pioneers act together to guide axons in the ventral cord of C. elegans. Development *130*, 5307-5318.

Hutter, H., Wacker, I., Schmid, C., and Hedgecock, E.M. (2005). Novel genes controlling ventral cord asymmetry and navigation of pioneer axons in C. elegans. Dev Biol *284*, 260-272.

Jacob, J., Kong, J., Moore, S., Milton, C., Sasai, N., Gonzalez-Quevedo, R., Terriente, J., Imayoshi, I., Kageyama, R., Wilkinson, D.G., *et al.* (2013). Retinoid acid specifies neuronal identity through graded expression of Ascl1. Curr Biol *23*, 412-418.

Jin, H., Emmons, S.W., and Kim, B. (2019). Expressional artifact caused by a co-injection marker rol-6 in C. elegans. PLoS One *14*, e0224533.

Johnson, D.S., Mortazavi, A., Myers, R.M., and Wold, B. (2007). Genome-wide mapping of in vivo protein-DNA interactions. Science *316*, 1497-1502.

Jolma, A., and Taipale, J. (2011). Methods for Analysis of Transcription Factor DNA-Binding Specificity In Vitro. Subcell Biochem *52*, 155-173.

Kerk, S.Y., Kratsios, P., Hart, M., Mourao, R., and Hobert, O. (2017). Diversification of C. elegans Motor Neuron Identity via Selective Effector Gene Repression. Neuron *93*, 80-98.

Koch, B., Nijmeijer, B., Kueblbeck, M., Cai, Y., Walther, N., and Ellenberg, J. (2018). Generation and validation of homozygous fluorescent knock-in cells using CRISPR-Cas9 genome editing. Nat Protoc *13*, 1465-1487.

Kratsios, P., Kerk, S.Y., Catela, C., Liang, J., Vidal, B., Bayer, E.A., Feng, W., De La Cruz, E.D., Croci, L., Consalez, G.G., *et al.* (2017). An intersectional gene regulatory strategy defines subclass diversity of C. elegans motor neurons. Elife *6*.

Kratsios, P., Pinan-Lucarre, B., Kerk, S.Y., Weinreb, A., Bessereau, J.L., and Hobert, O. (2015). Transcriptional coordination of synaptogenesis and neurotransmitter signaling. Curr Biol *25*, 1282-1295.

Lee, J.Y., and Goldstein, B. (2003). Mechanisms of cell positioning during C. elegans gastrulation. Development *130*, 307-320.

Letourneau, P.C., Condic, M.L., and Snow, D.M. (1992). Extracellular matrix and neurite outgrowth. Curr Opin Genet Dev *2*, 625-634.

Li, Q., Herrler, M., Landsberger, N., Kaludov, N., Ogryzko, V.V., Nakatani, Y., and Wolffe, A.P. (1998). Xenopus NF-Y pre-sets chromatin to potentiate p300 and acetylation-responsive transcription from the Xenopus hsp70 promoter in vivo. EMBO J *17*, 6300-6315.

Li, X.Y., Mantovani, R., Hooft van Huijsduijnen, R., Andre, I., Benoist, C., and Mathis, D. (1992). Evolutionary variation of the CCAAT-binding transcription factor NF-Y. Nucleic Acids Res *20*, 1087-1091.

Luo, R., Lu, J.F., Hu, Q., and Maity, S.N. (2008). CBF/NF-Y controls endoplasmic reticulum stress induced transcription through recruitment of both ATF6(N) and TBP. J Cell Biochem *104*, 1708-1723.

Maerkl, S.J., and Quake, S.R. (2007). A systems approach to measuring the binding energy landscapes of transcription factors. Science *315*, 233-237.

Mantovani, R. (1998). A survey of 178 NF-Y binding CCAAT boxes. Nucleic Acids Res *26*, 1135-1143.

Mantovani, R. (1999). The molecular biology of the CCAAT-binding factor NF-Y. Gene *239*, 15-27.

Masoudi, N., Tavazoie, S., Glenwinkel, L., Ryu, L., Kim, K., and Hobert, O. (2018). Unconventional function of an Achaete-Scute homolog as a terminal selector of nociceptive neuron identity. PLoS Biol *16*, e2004979.

Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.Y., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G., Worsley-Hunt, R.*, et al.* (2016). JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. Nucleic Acids Res *44*, D110-115.

Matys, V., Kel-Margoulis, O.V., Fricke, E., Liebich, I., Land, S., Barre-Dirrie, A., Reuter, I., Chekmenev, D., Krull, M., Hornischer, K.*, et al.* (2006). TRANSFAC and its module TRANSCompel: transcriptional gene regulation in eukaryotes. Nucleic Acids Res *34*, D108-110.

Mikuni, T., Nishiyama, J., Sun, Y., Kamasawa, N., and Yasuda, R. (2016). High-Throughput, High-Resolution Mapping of Protein Localization in Mammalian Brain by In Vivo Genome Editing. Cell *165*, 1803-1817.

Milton, A.C., Packard, A.V., Clary, L., and Okkema, P.G. (2013). The NF-Y complex negatively regulates Caenorhabditis elegans tbx-2 expression. Dev Biol *382*, 38-47.

Mitchell, P.J., and Tjian, R. (1989). Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins. Science *245*, 371-378.

Moody, S.A., Klein, S.L., Karpinski, B.A., Maynard, T.M., and LaMantia, A.S. (2013). On becoming neural: what the embryo can tell us about differentiating neural stem cells. American Journal of Stem Cells *2*, 74-94.

Moore, I., and Murphy, A. (2009). Validating the location of fluorescent protein fusions in the endomembrane system. Plant Cell *21*, 1632-1636.

Moresco, J.J., and Koelle, M.R. (2004). Activation of EGL-47, a Galpha(o)-coupled receptor, inhibits function of hermaphrodite-specific motor neurons to regulate Caenorhabditis elegans egg-laying behavior. J Neurosci *24*, 8522-8530.

Morey, M., Yee, S.K., Herman, T., Nern, A., Blanco, E., and Zipursky, S.L. (2008). Coordinate control of synaptic-layer specificity and rhodopsins in photoreceptor neurons. Nature *456*, 795-799.

Nardini, M., Gnesutta, N., Donati, G., Gatta, R., Forni, C., Fossati, A., Vonrhein, C., Moras, D., Romier, C., Bolognesi, M.*, et al.* (2013). Sequence-specific transcription factor NF-Y displays histone-like DNA binding and H2B-like ubiquitination. Cell *152*, 132-143.

Oldfield, A.J., Yang, P., Conway, A.E., Cinghu, S., Freudenberg, J.M., Yellaboina, S., and Jothi, R. (2014). Histone-fold domain protein NF-Y promotes chromatin accessibility for cell type-specific master transcription factors. Mol Cell *55*, 708-722.

Pacary, E., Heng, J., Azzarelli, R., Riou, P., Castro, D., Lebel-Potter, M., Parras, C., Bell, D.M., Ridley, A.J., Parsons, M.*, et al.* (2011). Proneural transcription factors regulate different steps of cortical neuron migration through Rnd-mediated inhibition of RhoA signaling. Neuron *69*, 1069-1084.

Pak, C.W., Flynn, K.C., and Bamburg, J.R. (2008). Actin-binding proteins take the reins in growth cones. Nat Rev Neurosci *9*, 136-147.

Peng, C.H., Hsu, J.T., Chung, Y.S., Lin, Y.J., Chow, W.Y., Hsu, D.F., and Tang, C.Y. (2006). Identification of degenerate motifs using position restricted selection and hybrid ranking combination. Nucleic Acids Res *34*, 6379-6391.

Pocock, R., and Hobert, O. (2008). Oxygen levels affect axon guidance and neuronal migration in Caenorhabditis elegans. Nat Neurosci *11*, 894-900.

Poon, G.M. (2010). Explicit formulation of titration models for isothermal titration calorimetry. Anal Biochem *400*, 229-236.

Prasov, L., Masud, T., Khaliq, S., Mehdi, S.Q., Abid, A., Oliver, E.R., Silva, E.D., Lewanda, A., Brodsky, M.C., Borchert, M.*, et al.* (2012). ATOH7 mutations cause autosomal recessive persistent hyperplasia of the primary vitreous. Hum Mol Genet *21*, 3681-3694.

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

Reiter, F., Wienerroither, S., and Stark, A. (2017). Combinatorial function of transcription factors and cofactors. Curr Opin Genet Dev *43*, 73-81.

Ritzefeld, M., and Sewald, N. (2012). Real-Time Analysis of Specific Protein-DNA Interactions with Surface Plasmon Resonance. J Amino Acids *2012*, 816032.

Rogers, C.D., Moody, S.A., and Casey, E.S. (2009). Neural induction and factors that stabilize a neural fate. Birth Defects Res C Embryo Today *87*, 249-262.

Romier, C., Cocchiarella, F., Mantovani, R., and Moras, D. (2003). The NF-YB/NF-YC structure gives insight into DNA binding and transcription regulation by CCAAT factor NF-Y. J Biol Chem *278*, 1336-1345.

Sandberg, M., Kallstrom, M., and Muhr, J. (2005). Sox21 promotes the progression of vertebrate neurogenesis. Nat Neurosci *8*, 995-1001.

Sasai, Y. (1998). Identifying the missing links: Genes that connect neural induction and primary neurogenesis in vertebrate embryos. Neuron *21*, 455-458.

Schmitz, A., and Galas, D.J. (1979). The interaction of RNA polymerase and lac repressor with the lac control region. Nucleic Acids Res *6*, 111-137.

Seeger, M., Tear, G., Ferres-Marco, D., and Goodman, C.S. (1993). Mutations affecting growth cone guidance in Drosophila: genes necessary for guidance toward or away from the midline. Neuron *10*, 409-426.

Serrano-Saiz, E., Leyva-Diaz, E., De La Cruz, E., and Hobert, O. (2018). BRN3-type POU Homeobox Genes Maintain the Identity of Mature Postmitotic Neurons in Nematodes and Mice. Curr Biol *28*, 2813-2823 e2812.

Shaham, S., and Bargmann, C.I. (2002). Control of neuronal subtype identity by the C. elegans ARID protein CFI-1. Genes Dev *16*, 972-983.

Sonnhammer, E.L., and Durbin, R. (1997). Analysis of protein domain families in Caenorhabditis elegans. Genomics *46*, 200-216.

Spitz, F., and Furlong, E.E. (2012). Transcription factors: from enhancer binding to developmental control. Nat Rev Genet *13*, 613-626.

Stefanakis, N., Carrera, I., and Hobert, O. (2015). Regulatory Logic of Pan-Neuronal Gene Expression in C. elegans. Neuron *87*, 733-750.

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.

Suzuki, Y., Tsunoda, T., Sese, J., Taira, H., Mizushima-Sugano, J., Hata, H., Ota, T., Isogai, T., Tanaka, T., Nakamura, Y.*, et al.* (2001). Identification and characterization of the potential promoter regions of 1031 kinds of human genes. Genome Res *11*, 677-684.

Tessier-Lavigne, M., and Goodman, C.S. (1996). The molecular biology of axon guidance. Science (New York, NY) *274*, 1123.

Tharyan, R.G., Annibal, A., Schiffer, I., Laboy, R., Atanassov, I., Weber, A.L., Gerisch, B., and Antebi, A. (2020). NFYB-1 regulates mitochondrial function and longevity via lysosomal prosaposin. Nature Metabolism *2*, 387-396.

Tjian, R., and Maniatis, T. (1994). Transcriptional activation: a complex puzzle with few easy pieces. Cell *77*, 5-8.

Tsalik, E.L., Niacaris, T., Wenick, A.S., Pau, K., Avery, L., and Hobert, O. (2003). LIM homeobox gene-dependent expression of biogenic amine receptors in restricted regions of the C. elegans nervous system. Dev Biol *263*, 81-102.

Uchikawa, M., Kamachi, Y., and Kondoh, H. (1999). Two distinct subgroups of Group B Sox genes for transcriptional activators and repressors: their expression during embryonic organogenesis of the chicken. Mechanisms of Development *84*, 103-120.

Vidal, B., Aghayeva, U., Sun, H., Wang, C., Glenwinkel, L., Bayer, E., and Hobert, O. (2017). An atlas of Caenorhabditis elegans chemoreceptor expression.

Vidal, B., Aghayeva, U., Sun, H., Wang, C., Glenwinkel, L., Bayer, E.A., and Hobert, O. (2018). An atlas of Caenorhabditis elegans chemoreceptor expression. PLoS Biol *16*, e2004218.

White, J.G., Southgate, E., Thomson, J.N., and Brenner, S. (1976). The structure of the ventral nerve cord of Caenorhabditis elegans. Philosophical transactions of the Royal Society of London Series B, Biological sciences *275*, 327-348.

White, J.G., Southgate, E., Thomson, J.N., and 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*, 1-340.

Yamanaka, T., Miyazaki, H., Oyama, F., Kurosawa, M., Washizu, C., Doi, H., and Nukina, N. (2008). Mutant Huntingtin reduces HSP70 expression through the sequestration of NF-Y transcription factor. EMBO J *27*, 827-839.

Yamanaka, T., Tosaki, A., Kurosawa, M., Matsumoto, G., Koike, M., Uchiyama, Y., Maity, S.N., Shimogori, T., Hattori, N., and Nukina, N. (2014). NF-Y inactivation causes atypical neurodegeneration characterized by ubiquitin and p62 accumulation and endoplasmic reticulum disorganization. Nat Commun *5*, 3354.

Yamanaka, T., Tosaki, A., Miyazaki, H., Kurosawa, M., Koike, M., Uchiyama, Y., Maity, S.N., Misawa, H., Takahashi, R., Shimogori, T.*, et al.* (2016). Differential roles of NF-Y transcription factor in ER chaperone expression and neuronal maintenance in the CNS. Sci Rep *6*, 34575.

Yoshioka, Y., Suyari, O., and Yamaguchi, M. (2008). Transcription factor NF-Y is involved in regulation of the JNK pathway during Drosophila thorax development. Genes Cells *13*, 117-130.

Zallen, J.A., Kirch, S.A., and Bargmann, C.I. (1999). Genes required for axon pathfinding and extension in the C. elegans nerve ring. Development *126*, 3679-3692.

Zemzoumi, K., Frontini, M., Bellorini, M., and Mantovani, R. (1999). NF-Y histone fold alpha1 helices help impart CCAAT specificity. J Mol Biol *286*, 327-337.

Zhang, G., Zheng, S., Liu, H., and Chen, P.R. (2015). Illuminating biological processes through site-specific protein labeling. Chem Soc Rev *44*, 3405-3417.

Zhao, C., and Emmons, S.W. (1995). A transcription factor controlling development of peripheral sense organs in C. elegans. Nature *373*, 74-78.

Zhao, X., Li, G., and Liang, S. (2013). Several affinity tags commonly used in chromatographic purification. J Anal Methods Chem *2013*, 581093.

Zhou, F.Q., and Cohan, C.S. (2004). How actin filaments and microtubules steer growth cones to their targets. J Neurobiol *58*, 84-91.

Zou, Y., Stoeckli, E., Chen, H., and Tessier-Lavigne, M. (2000). Squeezing axons out of the gray matter: a role for slit and semaphorin proteins from midline and ventral spinal cord. Cell *102*, 363-375.

8. Appendix

Supplemental Table 1. The table below represents all the fragments used on the *sra-6* promoter analysis experiments. Left to right, **Column 1–** primer name, **Column 2 –** primer sequence, **Column 3 –** Size of the promoter regions amplified, **Column 4 –** size of GFP sequence, **Column 5 –** 3'UTR present on the plasmid, **Column 6 –** total size of the amplicon, **Column 7 –** original plasmid name that was used to amplify the sample, **Column 8 –** strain number. Red and yellow squares represent strains that were not generated or numbering is N/A.

Primers	Sequence (Forward primer)	Promoter size (bp)	GFP (bp)	3'UTR (bp)	Total bp (*Psra-6* +GFP+UTR)	Plasmid	Strain
OPM3F	ggagtttgataatatctgacgc	2340	856	unc-54 (283)	4246	pPD95.75	RJP3650
OPM1F	tgtaagacatacgttagctgg	1966	856	unc-54 (283)	3872	pPD95.75	RJP3651
OPM2F	aagttttactcggccgatgc	1125	856	unc-54 (283)	3031	pPD95.75	RJP3722
OTR58	GTGAAAACCTCTGACACATGC	Reverse primer used with the forward primers stated above for the pPD95.75				pPD95.75	
Clonning of Full *sra-6* promoter		3775	927	let-858 (473)	5200	pPD117.01	RJP3726
OPM1F_XbaI	aaaaTCTAGAtgtaagacatacgttagctgg	862	927	let-858 (473)	2323	pPD117.01	RJP3723
OPM2R_SmaI	aaaacccgggGCATCGGCCGAGTAAAACTT						
OPM2.8F_XbaI	aaaaTCTAGAacgcaagtcagagagattcc	615	927	let-858 (473)	2066	pPD117.01	RJP3767
OPM2.85F_XbaI	aaaatctagaTAATCTAAAATTTTAATTGC	586	927	let-858 (473)	2037	pPD117.01	
OPM2.86F_XbaI	AAAATCTAGAtcacgcgtaaaggaagagac	541	927	let-858 (473)	1992	pPD117.01	RJP3768
OPM2.87F	caataaaagtacacctgacg	439	927	let-858 (473)	1895	pPD117.01	
OPM2.88F	aataagtgctcgttaagtaatg	359	927	let-858 (473)	1815	pPD117.01	
OPM2.89F	cccaatttttatttcagcgatttc	264	927	let-858 (473)	1720	pPD117.01	RJP4172
OPM2.90F	aggatttctagtttagtctttg	171	927	let-858 (473)	1627	pPD117.01	RJP4085
OPM2.91F	cgtttcatttaattgtaagg	84	927	let-858 (473)	1540	pPD117.01	RJP3969
OPM2.92F	catcaaacgccagttttctg	60	927	let-858 (473)	1511	pPD117.01	RJP3946
OPM2.94F	tttaattataagttttactcggc	34	927	let-858 (473)	1485	pPD117.01	
OPM2.93F	ttataagttttactcggccg	29	927	let-858 (473)	1480	pPD117.01	
OPM8R	GGGAAATGTGCGCGGAACC	Reverse primer used with the forward primers stated above for the pPD117.01				pPD117.01	

Supplemental Table 2. Alleles, transgenes and reporter strains generated in chapter 4. Left – strain number or transgene reference, **Middle –** description of mutation and transgenes present in each strain, **Right –** outcross number. N/A – Not available.

Strains number	**Description**	**Outcross number**
N/A	*ceh-43(ot406) III; vtIs1 V (dat-1p::GFP + rol-6(su1006)]; oyIs14;*	OC1x
N/A	*cfi-1(ky651) I; oyIs14*	OC1x
Strains number	**Description**	**Outcross number**
RJP3862	*stIs11350 [Pnfya-1::H1-wCherry]; oyIs14 V*	OC1x
RJP4204	*stIs10879 [Pnfyb-1::H1-wCherry]; oyIs14 V*	OC1x
RJP3823	*stIs11751 [Pnfyc-1::H1-wCherry]; oyIs14 V*	OC1x
RJP3889	*nfya-1(ok1174) X; oyIs14 V*	OC4x
RJP3966	*nfya-1(ok1174) X; kyIs321*	OC4x

RJP3890	*nfya-1(ok1174) X; oyIs14 V; nfya-1 fosmid (WRM0633cC02)(1ng/ul); Pmyo-2::RFP(3ng/ul); line 1*	OC3x
RJP3891	*nfya-1(ok1174) X; oyIs14 V; nfya-1 fosmid (WRM0633cC02) (1ng/ul); Pmyo-2::RFP(3ng/ul) line 2*	OC3x
RJP4016	*nfya-1(ok1174) X; Pnpr-11::nfya-1 (cDNA) (1ng/ul); Pmyo-2::RFP(3ng/ul); Line 1*	OC3x
RJP4053	*nfya-1(ok1174) X; Pnpr-11::nfya-1 (cDNA) (1ng/ul); Pmyo-2::RFP (3ng/ul); Line 2*	OC3x
RJP3924	*nfya-1(ok1174) X; hdIs26*	OC4x
RJP3925	*kyIs321; mjIs27;*	-
RJP3926	*nfya-1(ok1174) X; kyIs321; mjIs27*	OC5x
RJP3925	*kyIs321; otEx6860*	-
RJP3927	*nfya-1(ok1174) X; kyIs321; otEx6860*	OC5x
RJP4086	*kyIs321; otEx6666*	OC4x
RJP4087	*nfya-1(ok1174) X; otEx6666*	OC5x
RJP4205	*kyIs321; otEx233*	OC1x
RJP4249	*nfya-1(ok1174) X; kyIs321; otEx233*	OC5x
RJP3862	*nfya-1(ok1174) X; oyIs14 V; stIs1135*	OC5x
RJP4130	*kyIs321; otEx6403*	-
RJP4129	*nfya-1(ok1174) X; kyIs321; otEx6403*	OC5x
RJP4374	*jvsEx428; kyIs321;*	-
RJP4375	*nfya-1 (ok1174) X; jvsEx428; kyIs321*	OC5x
RJP4526	*rpEx2016; kyIs321;*	-
RJP4541	*nfya-1 (ok1174) X; rpEx2016; kyIs321*	OC5x
RJP4531	*rpEx2017;*	-
RJP4532	*nfya-1(ok1174) X; rpEx2017*	OC5x
RJP4522	*otIs339; kyIs321*	-
RJP4524	*nfya-1 (ok1174) X; otIs339; kyIs321*	OC5x
RJP4439	*otIs11; kyIs321*	OC4x
RJP4525	*nfya-1 (ok1174) X; otIs11; kyIs321*	OC5x
RJP4533	*nfyc-1 (rp152) II; nfyb-1 (Cu13) II; oyIs14*	-
RJP4647	*nfya-1 (ok1174) X; nfya-2 (tm4194) I; oyIs14*	OC4x
RJP4646	*nfya-1 (ok1174) X; nfya-2 (tm4194) I; oyIs14; kyIs321*	OC4x
RJP4627	*stIs11350 [Pnfya-1::H1-wCherry]; otIs287 [rab-3(prom1)::2xNLS::YFP] IV*	-
RJP4628	*nfya-1 (ok1174) X; stIs11350 [Pnfya-1::H1-wCherry]; otIs287 [rab-3(prom1)::2xNLS::YFP] IV*	OC5x
RJP4720	*otIs619 (unc-11(prom8)::2xNLS::TagRFP)*	OC1x
RJP4721	*nfya-1 (ok1174) X; otIs619 [unc-11(prom8)::2xNLS::TagRFP]*	OC4x
RJP4753	*evIs82b [Punc-129::GFP + dpy-20(+)] IV*	OC1x
RJP4754	*evIs82b [Punc-129::GFP + dpy-20(+)] IV; stIs11350 [Pnfya-1::H1-wCherry]*	OC1x
RJP4755	*nfya-1 (ok1174) X; evIs82b [Punc-129::GFP + dpy-20(+)] IV*	OC5x

RJP4750	*otIs350 [ric-4(fosmid)::SL2::NLS::YFP::H2B + pha-1(+)]*	OC1x
RJP4751	*otIs350 [ric-4(fosmid)::SL2::NLS::YFP::H2B + pha-1(+)]; stIs11350 [Pnfya-1::H1-wCherry]*	OC1x
RJP4752	*nfya-1 (ok1174) X; otIs350 [ric-4(fosmid)::SL2::NLS::YFP::H2B + pha-1(+)]*	OC5x
RJP4765	*rpEx2121 (Psrg-32::GFP) (50ng/ul); pmyo-2::RFP* (3ng/ul)(Line 1) ; *kyIs321 (Pnpr-11::RFP)*	OC1x
RJP4766	*nfya-1 (ok1174)X; rpEx2121 (Psrg-32::GFP) (50ng/ul); pmyo-2::RFP* (3ng/ul)(Line 1) ; *kyIs321 (Pnpr-11::RFP)*	OC5x
RJP4498	*rpEx2006 [Psri-1::GFP(50ng/ul); pmyo-2::RFP(3ng/ul)]* WT line 1	-
RJP4499	*rpEx2007 [Psri-1::GFP(50ng/ul); pmyo-2::RFP(3ng/ul)]* WT line 2	-
RJP4500	*rpEx2008 [Psri-1::GFP(50ng/ul); pmyo-2::RFP(3ng/ul)] Mutagenesis* line 1	-
RJP4501	*rpEx2009 [Psri-1::GFP(50ng/ul); pmyo-2::RFP(3ng/ul)] Mutagenesis* line 2	-
RJP4502	*rpEx2010 [Psri-1::GFP(50ng/ul); pmyo-2::RFP(3ng/ul)] Mutagenesis* line 3	-
RJP4604	*rpEx2061 [Psrh-277::GFP(50ng/ul); pmyo-2::RFP(3ng/ul)]* WT line 1	-
RJP4618	*rpEx2071[Psrh-277::GFP(50ng/ul); pmyo-2::RFP(3ng/ul)]* WT line 2	-
RJP4605	*rpEx2062 [Psrh-277::GFP(50ng/ul); pmyo-2::RFP(3ng/ul)] Mutagenesis* line 1	-
RJP4606	*rpEx2063 [Psrh-277::GFP (50ng/ul); pmyo-2::RFP(3ng/ul)] Mutagenesis* line 2	-
RJP4763	*rpEx2121 [Psrg-32::GFP (50ng/ul); pmyo-2::RFP(3ng/ul)] WT line 1*	-
rpEx2017	*otIs287 [rab-3(prom1)::2xNLS::YFP] IV; Pnpr-11::WrmScarlet (10ng/ul)*	-
RJP4052	*nfya-1(bp4) X; oyIs14 V*	OC3x
RJP3812	*nfyb-1(cu13) II; oyIs14 V*	OC3x
RJP3965	*nfyb-1(cu13) II; kyIs321*	OC4x
RJP3967	*nfyb-1(tm4257) II; oyIs14 V*	OC1x
RJP3944	*nfyc-1(tm4145) II; oyIs14 V*	OC3x
RJP4034	*nfyc-1(tm4145) II; kyIs321*	OC4x
RJP4043	*nfyc-1(rp120) II; oyIs14 V*	OC1x
RJP4008	*nfya-2(tm4194) I; oyIs14 V*	OC3x
RJP4533	*rp152; nfyb-1 (cu13) II; oyIs14 V*	-

***S*upplemental Table 3. Alleles, transgenes and reporter strains generated in chapter 5. Left –** strain number or transgene reference, **Middle –** description of mutation and transgenes present in each strain, **Right –** outcross number.

Number	Description	outcross number
ROP01	V5::NFYA-1; *PHX2237(syb2237)*	OC0x
RJP4645	V5::NFYA-1; *PHX2237(syb2237)*; *oyIs14*	OC2x
RJP4735	V5::NFYA-1; *PHX2237(syb2237);*	OC3x
rp132	HA::NFYB-1	OC0x
RJP4427	HA::NFYB-1; *oyIs14*	OC2x
RJP4496	HA::NFYB-1	OC3x
rp130	HA::NFYC-1	OC0x
RJP4428	HA::NFYC-1; *oyIs14*	OC2x
RJP4497	HA::NFYC-1	OC3x