2010-07-30

C. Elegans Metabolic Gene Regulatory Networks: A Dissertation

H. Efsun Arda
University of Massachusetts Medical School

C. ELEGANS METABOLIC GENE REGULATORY NETWORKS

A Dissertation Presented

By

HATİCE EFSUN ARDA

Submitted to the Faculty of the

University of Massachusetts Graduate School of Biomedical Sciences, Worcester

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

JULY 30, 2010

INTERDISCIPLINARY GRADUATE PROGRAM

C. ELEGANS METABOLIC GENE REGULATORY NETWORKS

A Dissertation Presented By

Hatice Efsun Arda

The signatures of the Dissertation Defense Committee signifies completion and approval as to style and content of the Dissertation

_____

A.J. Marian Walhout, Ph.D., Thesis Advisor

_____

Eric Baehrecke, Ph.D., Member of Committee

_____

Job Dekker, Ph.D., Member of Committee

_____

Craig Mello, Ph.D., Member of Committee

_____

Marc Vidal, Ph.D., Member of Committee

The signature of the Chair of the Committee signifies that the written dissertation meets the requirements of the Dissertation Committee

_____

Heidi Tissenbaum, Ph.D., Chair of Committee

The signature of the Dean of the Graduate School of Biomedical Sciences signifies that the student has met all graduation requirements of the School

_____

Anthony Carruthers, Ph.D.
Dean of the Graduate School of Biomedical Sciences

Interdisciplinary Graduate Program

July 30, 2010

COPYRIGHT INFORMATION

DEDICATION

To my dear family and my beloved friends,

and also to those exceptional teachers and professors who selflessly devote themselves to educate and set young minds free.

Doktora tezim,

biricik aileme ve sevgili dostlarıma,

ve ayrıca kendilerini genç nesillerin eğitimine adayıp, onları özgür kılan tüm değerli öğretmenlere armağan olsun.

ACKNOWLEDGEMENTS

Firstly, I would like to thank my advisor, Dr. Marian Walhout for introducing me to the exciting field of systems biology, which I had been seeking to find and did not know if it existed until I joined the lab. Since the early days of my thesis work, Marian had always paid close attention to the progress of my projects; encouraging me to keep going when things did not work out, and appreciating with enthusiasm when there was success. From designing a controlled experiment to making figures for a manuscript, I am forever grateful for her efforts in training me as an independent scientist.

I thank my former colleagues, Bart Deplancke, Vanessa Vermeirssen, Natalia Martinez and Christian Grove for the wonderful atmosphere that their presence created in the lab, which was often filled with loud music, lots of laughter, and heated discussions about science and pretty much everything else. I would also like to thank the present members of the Walhout lab, particularly John Reece-Hoyes and Lesley MacNeil for their friendship and willingness to help on matters both scientific and personal. In addition, I thank the past and present members of the Dekker, Tissenbaum and Mello labs, who were always available to share scientific expertise and technical assistance. It was a real pleasure to be a part of this community.

I thank the members of my thesis advisory committee, Drs. Michael Czech, Craig Peterson, Yong-Xu Wang, Craig Mello and Heidi Tissenbaum for their continuous support and guidance throughout my thesis work. I also thank Drs. Job Dekker, and Eric Baehrecke for making the time and agreeing to be a member on my dissertation committee. I am especially grateful to Dr. Marc Vidal for accepting to be my external committee member.

I would like to express my gratitude to the Graduate School of Biomedical Sciences staff, Program in Gene Function and Expression faculty, lab members and staff, especially to Nina Bhabhalina, Sharon Briggs, Elizabeth Nourse and Randi Wasik for their technical and administrative support.

My final acknowledgments go to my dear friends, here and abroad, who enrich my life with their love and company, and make me feel home when it is not easy to be a graduate student abroad; to my dear family, who always supported my unconventional career choices with endless patience and understood that I will forever stay a student of science and never get a ‘real’ job. Also, to my special friends Minik and Sinzu, who bring joy and happiness to my life. Last, but not least, I thank Pedro Batista who has been an incredible friend, life partner, colleague and family, and his giant soul and wisdom continues to enlighten me.

ABSTRACT

In multicellular organisms, determining when and where genes will be expressed is critical for their development and physiology. Transcription factors (TFs) are major specifiers of differential gene expression. By establishing physical contacts with the regulatory elements of their target genes, TFs often determine whether the target genes will be expressed or not. These physical and/or regulatory TF-DNA interactions can be modeled into gene regulatory networks (GRNs), which provide a systems-level view of differential gene expression. Thus far, much of the GRN delineation efforts focused on metazoan development, whereas the organization of GRNs that pertain to systems physiology remains mostly unexplored.

My work has focused on delineating the first gene regulatory network of the nematode *Caenorhabditis elegans* metabolic genes, and investigating how this network relates to the energy homeostasis of the nematode. The resulting metabolic GRN consists of ~70 metabolic genes, 100 TFs and more than 500 protein–DNA interactions. It also includes novel protein-protein interactions involving the metabolic transcriptional cofactor MDT-15 and several TFs that occur in the metabolic GRN. On a global level, we found that the metabolic GRN is enriched for nuclear hormone receptors (NHRs). NHRs form a special class of TFs that can interact with diffusible biomolecules and are well-known regulators

of lipid metabolism in other organisms, including humans. Interestingly, NHRs comprise the largest family of TFs in nematodes; the *C. elegans* genome encodes 284 NHRs, most of which are uncharacterized. In our study, we show that the *C. elegans* NHRs that we retrieved in the metabolic GRN organize into network modules, and that most of these NHRs function to maintain lipid homeostasis in the nematode. Network modularity has been proposed to facilitate rapid and robust changes in gene expression. Our results suggest that the *C. elegans* metabolic GRN may have evolved by combining NHR family expansion with the specific modular wiring of NHRs to enable the rapid adaptation of the animal to different environmental cues.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

Preface to Chapter I

This chapter introduces the concept and significance of gene regulatory networks (GRNs), and those that pertain to systems physiology. Also introduced is the use of *Caenorhabditis elegans* as a model system to study physiological GRNs, and the role of transcriptional regulation on energy homeostasis with the emphasis on nuclear hormone receptors.

Most of this chapter has been published separately in:

Arda, H. E., and Walhout, A. J. (2009). Gene-centered regulatory networks. Brief Funct Genomic Proteomic *9*, 4-12.

Arda, H. E., Taubert, S., MacNeil, L. T., Conine, C. C., Tsuda, B., Van Gilst, M., Sequerra, R., Doucette-Stamm, L., Yamamoto, K. R., and Walhout, A. J. M. (2010). Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. Mol. Syst. Biol *6*, 367.

CHAPTER I

Gene-centered Regulatory Networks

Introduction

Differential gene expression is essential for the establishment of body plans and physiology of multicellular organisms. Genes need to be turned on or off at the right place and the right time in order to instruct organogenesis and to maintain post-developmental physiology. In the developing mouse embryo, Oct4, Sox2, Klf4 and c-Myc genes are expressed early during development and function to preserve the pluripotency of uncommitted stem cells (Takahashi et al., 2007). Their expression is sufficient to preserve or induce this unique cellular property, while their misexpression later in life can cause cancer. On the other hand, a different group of genes is activated upon cellular differentiation, such as the Pax6 gene that is required for eye development in a variety of organisms (Gehring and Ikeo, 1999). Likewise, in *Drosophila* embryos, the spatiotemporal expression of *gap* and *pair-rule* genes is crucial for defining segmentation patterns and development (St Johnston and Nüsslein-Volhard, 1992). Finally, yet another group of genes is expressed in specialized, fully differentiated tissues to enable post-developmental functions in physiology throughout the lifetime of an organism. A textbook example is the insulin precursor gene, which is specifically expressed in the pancreas and regulates the level of glucose in the blood after food intake.

Differential gene expression is a highly regulated and controlled process. It occurs at the first level by the action of regulatory transcription factors (TFs):

proteins that recognize specific nucleotide sequences and initiate contacts with genomic DNA at regions containing these sequences. These sites on genomic DNA are referred to as '*cis*-regulatory elements', and are usually located in the promoter or enhancer regions of target genes [reviewed in (Walhout, 2006)]. The interactions between TFs and *cis*-regulatory elements often result in either activation or repression of target gene transcription, depending on the cellular context. In addition to TFs, transcriptional cofactors, proteins that interact with or are part of the core transcriptional machinery or protein complexes that modify chromatin (*e.g.* histone acetylation, methylation, *etc*.) can modulate the function of TFs by forming physical interactions with TFs (Näär et al., 2001; Rando and Chang, 2009). Also, RNA binding proteins, mRNA stability, export and splicing, microRNAs [reviewed in: (Ambros, 2004)], and post-translational modifications all contribute to differential gene expression. However, it is transcriptional regulation which first and foremost determines where and when a gene is expressed, whereas other types of regulation often modulate and dampen gene expression, rather than to determine it.

What are gene regulatory networks?

Physical and regulatory interactions between genes and their regulators, as well as the interactions within the regulators can be visualized as graph models. These models often have the form of a network, and are collectively referred to as *gene regulatory networks* (GRNs) (Erwin and Davidson, 2009). In

these networks, the regulatory components and targets are the ‘nodes’, and the relationships between them are the ‘edges’ of the network (Walhout, 2006; Arda and Walhout, 2009) (Figure 1.1). The availability of genome sequences of diverse organisms and the development of genome-scale, high-throughput technologies allowed researchers to collect large amounts of information regarding differential gene expression. The GRN models generated using high-throughput data reveals recurring patterns of gene regulation on a global scale across different species (see below), and provides novel insights into the mechanisms of transcriptional regulation by connecting global and local network organization to network functionality (Milo et al., 2002; Segal et al., 2003; Babu et al., 2004; Vermeirssen et al., 2007a; Martinez et al., 2008).

GRNs are often highly complex, particularly when large numbers of interactions are incorporated, which makes them difficult to interrogate by eye (Figure 1.2A). Instead, a variety of computational and mathematical tools for network analysis need to be used. These tools can inform us about the properties of the network as a whole, or identify important network neighborhoods such as 'modules', or find overrepresented network building blocks, such as 'motifs' (Figure 1.2B). At the level of whole networks, measures that are often used are the degree and degree distribution (Barabási and Oltvai, 2004). The degree is defined as the connectivity of individual nodes, *i.e.* the number of TFs bound by a DNA fragment or the number of DNA fragments bound by a TF. These are referred to as incoming and outgoing degree,

respectively. The degree distribution provides information about the overall network connectivity. In the majority of biological networks, the degree distribution follows a power law, rather than a normal distribution: most nodes have a relatively low degree, but a small number are disproportionally highly connected (Barabasi and Albert, 1999). These TFs and promoters are referred to as ‘hubs’. TF hubs interact with many genes, from many different tissues or organs. They are often essential, indicating their overall importance in gene regulation and development (Yu et al., 2004; Deplancke et al., 2006).

‘TF modules’ are highly interconnected network neighborhoods that occur in GRNs. Such modules can contain functionally related genes and TFs. Several measures can be used to identify TF modules, including topological overlap coefficient (TOC) analysis, which essentially determines the similarity between TFs based on the target genes they share in the network (Ravasz et al., 2002; Vermeirssen et al., 2007a). Each pair of TFs is given a similarity or a TOC score. These scores are then used to generate a hierarchical clustering dendogram in which TFs with similar connections are positioned next to each other and this allows the identification of TF modules (Figure 1.2B). Vermeirssen *et al*. previously identified TF modules in a neuronal GRN and found that one of these was enriched for paired-type homeodomain TFs (Vermeirssen et al., 2007a). Interestingly, these TFs associate predominantly with genes that are exclusively neuronal and these TFs are themselves neuronally expressed. Moreover, the neuronal expression of paired-type homedomain TFs is conserved between *C.*

elegans and mice. This example illustrates how TF modules can be used to functionally annotate sets of TFs at a systems level.

Network motifs are small building blocks composed of two or more interactions that are overrepresented in GRNs compared to randomized networks (Milo et al., 2002). Thus, such motifs represent successful mechanisms of gene regulation. One of the best-studied motifs is the feed-forward loop in which a TF regulates another TF and both share a downstream target (Mangan and Alon, 2003) (Figure 1.2B). This motif is found in GRNs from bacteria, yeast and *C. elegans* (Deplancke et al., 2006; Milo et al., 2002). Martinez *et al*. recently delineated the first genome-scale GRN that contains interactions between microRNA promoters and TFs, and predicted interactions between microRNAs and the 3'UTRs of TF-encoding mRNAs. Interestingly, this network contains a novel type of network motif: a feedback loop in which microRNAs and TFs reciprocally regulate each other (Figure 1.2B). The microRNAs and TFs that participate in these loops have a high *flux-capacity:* a combined high in-and outgoing degree. This indicates that such motifs may provide adaptable and robust gene expression programs [reviewed in (Martinez and Walhout, 2009)].

Experimental approaches to map GRNs: TF-centered vs. gene-centered methods

Physical TF-DNA interactions can be experimentally delineated using two conceptually different but highly complementary approaches (Figure 1.3 and

Table 1.1). TF-centered (protein-to-DNA) methods start with a TF or set of TFs of interest and identify genomic DNA fragments with which these TF(s) interact. Chromatin immunoprecipitation (ChIP), protein binding microarrays (PBMs), and DamID are the most widely used TF-centered methods (van Steensel et al., 2001; Mukherjee et al., 2004; Kim and Ren, 2006). ChIP has been particularly powerful for the identification of TF-DNA interactions in homogeneous systems such as yeast, and in mammalian tissue culture cells, including primary cells or stem cells. Although powerful, it is technically difficult to systematically apply ChIP to most TFs in heterogeneous and complex metazoan systems such as intact worms. Especially if the TFs are expressed at low levels, and some TFs may be expressed in a limited number of cells, or during a narrow developmental interval, therefore complicating the detection of specific signals. Further, the analysis of TF-DNA interactions in complex systems has been limited to a handful of TFs, due to the availability of antibodies that are suitable for ChIP assays. Therefore, complimentary approaches are needed for the delineation of GRNs.

Gene-centered (DNA-to-protein) methods start with one or more regulatory DNA fragments and identify the TFs that can interact with these fragments (Walhout, 2006; Arda and Walhout, 2009). Seminal work from Eric Davidson and colleagues, who characterized the wiring of endo-mesodermal gene regulation in the sea urchin embryo, epitomizes the concept of gene-centered regulatory network mapping (Davidson et al., 2002). This work focused

on the identification of *cis*-regulatory DNA elements to understand where and when genes are expressed during development, followed by the discovery of TFs that may regulate this process. The endo-mesodermal network has been delineated over the course of many years because the work required laborious assays that were not amenable to use in high-throughput settings in complex animals. In addition, the interactions identified are not necessarily direct, as physical associations are not immediately revealed. To provide a gene-centered method that can detect physical interactions between sets of genes and multiple TFs in a relatively short amount of time, we have developed a high-throughput version of the yeast one-hybrid (Y1H) system (Deplancke et al., 2004). With this method, interactions identified are strictly physical, as Y1H assays do not immediately reveal *in vivo* regulatory consequences (see also below).

Table 1.1 Advantages and limitations of TF-centered *versus* gene-centered methods

TF-centered method (ChIP)	Gene-centered method (Y1H)
Start with a TF and find all possible genomic regions the TF can bind	Start with a genomic fragment (*e.g.* a gene promoter) and find TFs that can interact with the DNA fragment
Interactions captured *in vivo*	Interactions captured are '*in levuro'*
Powerful for the identification of protein-DNA interactions in unicellular or homogenous systems	Powerful for the identification of protein-DNA interactions that pertain to complex tissues or processes
Powerful for capturing interactions when TF activity depends on post-transcriptional modifications or TF functions as a heterodimer	May miss interactions that involve TFs that are post-transcriptionally modified, and cannot yet detect heterodimers
Limited by the endogenous expression level and/or spatio-temporal expression pattern of TFs, *i.e.* TFs that are expressed at high levels, or in a large number of cells can be assayed	Not limited by the endogenous expression level and/or spatio-temporal state of TFs, interactions found are 'condition-independent'
Limited by the availability of TF-specific antibodies	No antibody is required

Gateway-compatible yeast one-hybrid assays

Y1H assays are conceptually similar to yeast two-hybrid (Y2H) assays, which have been used to identify thousands of protein-protein interactions in many systems, including *C. elegans* (Walhout et al., 2000; Li et al., 2004). Instead of using two hybrid proteins (a protein bait and a protein prey), the Y1H system uses a DNA bait and a single hybrid protein prey (Figure 1.4). The Y1H system was first developed to facilitate the identification of proteins that can bind to multiple copies of a short DNA sequence of interest (Li and Herskowitz, 1993). However, the comprehensive mapping of GRNs needs to be unbiased, as most small *cis*-regulatory DNA elements that control gene expression are not yet discovered. Instead, larger genomic fragments that likely harbor many of these elements such as gene promoters and enhancers need to be interrogated. Further, the original Y1H system used conventional, restriction-enzyme-based cloning methods for DNA bait generation and was therefore not amenable for system-level analyses. To alleviate these limitations, we have developed a Y1H system that uses Gateway recombinational cloning to rapidly transfer multiple DNA baits into Y1H Destination vectors in parallel (Deplancke et al., 2004). We have demonstrated that this Y1H system can be used with both small elements and with large, complex DNA fragments such as gene promoters.

Gateway-compatible Y1H assays start with a set of DNA fragments (DNA baits) of interest (Figure 1.4). Briefly, the DNA bait is Gateway-cloned upstream of two Y1H reporter genes (*HIS3* and *LacZ*) and the two DNA bait::reporter

constructs are integrated into the yeast genome. This ensures that DNA baits are chromatinized which minimizes background and, therefore, reduces false positive interactions. To enable the identification of a wide variety of DNA binding proteins, including transcriptional repressors, a strong heterologous activation domain (AD – Figure 1.4) is added to the prey proteins. If the prey protein contains a DNA binding domain that can interact with the DNA bait, reporter gene expression is activated. Activation of *HIS3* expression is assessed on media lacking histidine and containing 3-aminotriazole (3-AT), a competitive inhibitor of the His3 enzyme. Activation of *LacZ* is assessed by a colorimetric ('blue-white') assay (Figure 1.4). So far, we have used both a cDNA library and TF library as prey resources in our Y1H screens. It was important to include a low-complexity TF library because TFs that are expressed at low levels or in only a few cells in an organism are difficult to retrieve from screens that employ high-complexity, non-normalized cDNA libraries. One can also perform mating or transformation experiments with arrays of TFs to directly test TF-DNA interactions (an example of five TFs is shown in Figure 1.4) (Vermeirssen et al., 2007b). In addition to predicted TFs, we have also retrieved several proteins that do not possess a recognizable DNA binding domain. These proteins robustly interact with their Y1H targets in yeast, and for 11 of them we have confirmed direct promoter interactions by ChIP from yeast using an anti-AD antibody (Deplancke et al., 2006; Vermeirssen et al., 2007a).

GRNs pertaining to systems physiology

Thus far, much focus has been on mapping GRNs related to animal development and understanding the establishment of body plans of multicellular organisms; even though differential gene expression continues to play an important role after embryonic development (Harbison et al., 2004; Levine and Davidson, 2005; Resendis-Antonio et al., 2005; Sandmann et al., 2007). Indeed, numerous human diseases, such as obesity, diabetes and cancer are characterized by profound changes in gene expression.

The delineation of GRNs that pertain to systems physiology has been limited to those of unicellular organisms such as *E.coli* and yeast (Ihmels et al., 2004; Barrett et al., 2005; Salgado et al., 2006) but has been a grand challenge for multicellular organisms. One of the reasons is that maintenance of energy homeostasis is an extremely complex process, and involves multiple physiological systems that function at the level of the whole organism. In vertebrates for instance, substantial crosstalk exists between the sites of energy storage and utilization (adipose and muscle systems) that gauge the energy state of the organism and the sites that modulate nutrient seeking behavior (nervous and endocrine systems).

C. elegans *as a model system to study GRNs pertaining to systems physiology*

Systems level studies of differential gene expression are greatly advanced by the use of genetically tractable model organisms such as the nematode

Caenorhabditis elegans. Below is an overview of features that make *C. elegans* a powerful model organism to study GRNs pertaining to systems physiology.

Functional genomics aspect: *C. elegans* is a free-living nematode with a relatively simple body plan and a fixed lineage of 959 somatic cells (Figure 1.5A). Under laboratory conditions, lifecycle can be completed approximately in three days. Once hatched from an egg, the animal goes through four larval molts before it reaches adulthood. A single adult hermaphrodite can produce ~300 progeny in its lifetime. Hermaphrodites (XX) are the predominant sex in a *C. elegans* population, while male *C. elegans* (XO) can occur with ~0.1% frequency due to sex chromosome non-disjunction in the hermaphrodite germline (Wood, 1988). The short lifecycle and the existence of both sexes enabled *C. elegans* researchers to perform numerous genetic screens, isolate many mutants and generate genetic tools to study gene function and regulation.

The *C. elegans* genome is fully sequenced and annotated, and is compact compared to the human genome: even though both contain ~20,000 genes, the 100 Mb worm genome is 30 times smaller (International Human Genome Sequencing Consortium, 2004; *C. elegans* Sequencing Consortium, 1998). Consequently, ~26% of the worm genome is exonic, compared to 1-2% in humans. In addition, the majority of intergenic regions are shorter than 2 kb (Dupuy et al., 2004), and introns are much shorter with a median length of 65 bp, whereas the median length of human introns is 3 kb (Spieth and Lawson, 2006). Thus, the potential regulatory genomic ‘space’ that needs to be considered in

studies of differential gene expression is much smaller. The *C. elegans* genome also encodes fewer TFs (~940) than the human genome (~1,500) (Reece-Hoyes et al., 2005; Vermeirssen et al., 2007b; Vaquerizas et al., 2009). There are also several invaluable 'omic' resources, such as the 'Orfeome', a collection of ~12,500 verified *C. elegans* open reading frames, the 'Promoterome', which contains ~5,000 *C. elegans* promoters, and the RNA interference (RNAi) library with 16,000 clones for reverse genetics studies (Lamesch et al., 2004; Dupuy et al., 2004; Kamath et al., 2003). Finally, studies about the mechanisms of differential gene expression at a systems level are greatly facilitated by the fact that *C. elegans* is a transparent animal. By using reporters such as the green fluorescent protein (GFP) one can elucidate where and when genes are expressed in living animals, and determine how different perturbations affect gene expression (Chalfie et al., 1994; Reece-Hoyes et al., 2007; Martinez et al., 2008)

Physiological aspect: *C. elegans* is a cholesterol auxotroph and feeds on *E. coli*, although its complete dietary range is yet to be determined. In the laboratory, it can be cultivated on either solid agar plates or in liquid medium supplemented with cholesterol and its bacterial diet (Brenner, 1974). The digestive tract of *C. elegans* consists of the 'pharynx', through which the bacteria are ingested and mechanically disrupted by the pharyngeal pumping action and passed to the intestine (Figure 1.5B). The intestine is a multifunctional organ that is responsible for the digestion and absorption of nutrients, synthesizing

macromolecules, clearing toxins, and defending against pathogens by initiating innate immune response (McGhee, 2007).

Importantly, *C. elegans* maintains its energy balance through diverse physiological systems that include neuroendocrine signaling, storing, mobilizing and utilizing energy stores (Ashrafi, 2007). *C. elegans* does not have specialized adipocytes, however it is capable of storing its body fat mainly in the form of triglycerides, which is similar to the mammalian body fat (Mullaney and Ashrafi, 2009). Whole animal staining assays using lipophilic dyes revealed that *C. elegans* body fat is mostly stored in the intestinal and skin-like hypodermal cells (Figure 1.5C). Despite the lack of extensive compartmentalization and tissue specialization that exist in mammalian metabolic systems, many pathways that are essential for mammalian metabolism can be found in *C. elegans* to some degree; for instance, the components of the insulin-like signaling, the serotonin signaling, target of rapamycin signaling and enzymes that function in core metabolic pathways including mitochondrial β-oxidation and fatty acid synthesis, elongation and desaturation (Ogg et al., 1997; Srinivasan et al., 2008; Soukas et al., 2009; Van Gilst et al., 2005b; Watts and Browse, 2002).

The ease of forward and reverse genetic screens, and the availability of molecular tools to dissect lipid metabolism enabled *C. elegans* researchers to identify hundreds of genes in less than a decade of time. By combining genome wide RNAi screening with vital lipid staining, Ashrafi *et al.* identified ~400 genes that when knocked down, alter the intensity of the staining of wild-type animals,

which encode for enzymes that function in lipid synthesis or breakdown pathways, membrane receptors, vesicle transporters, and notably TFs (Ashrafi et al., 2003). In addition to the effects of gene inactivations observed in wild-type animals, when the authors tested these 400 genes by RNAi knockdown in different neuroendocrine mutant strains, including *daf-2* (insulin-like signaling receptor mutant)*, tub-1* (tubby ortholog) and *tph-1* (serotonin pathway mutant), they found that 32 of them caused reduced staining in all mutant animals. These genes were referred as 'core fat genes', and based on their RNAi phenotype across diverse neuroendocrine mutant strains; the authors suggested that these genes might be key to lipid homeostasis.

C. elegans is also equipped with systems necessary for nutrient sensing and starvation response. Van Gilst *et al.* surveyed the transcript levels of an extensive list of *C. elegans* glucose and lipid metabolism genes under well fed and short term fasting conditions (Van Gilst et al., 2005b). The authors found that expression of several genes that function in mitochondrial and peroxisomal β-oxidation, fatty acid desaturation, and fatty acid binding and transport were markedly altered as the animals were switched between the two nutrient states. Due to the observed transcriptional response the authors referred these genes as 'fasting response genes'.

The role of transcriptional regulation in C. elegans *lipid homeostasis*

The finding that *C. elegans* actively adjusts the expression of its metabolic genes during starvation led to the hypothesis that dynamic transcriptional programs must exist, which integrate external and internal nutrient signals to accommodate the energy needs of the animal. Remarkably, several TFs that have been characterized in mammalian systems that function in lipid homeostasis are found to be orthologous in *C. elegans* (Ashrafi, 2007). For instance, the sterol regulatory element binding protein (SREBP) in mammals regulates cholesterol production and fatty acid synthesis (Brown and Goldstein, 1997). Although *C. elegans* does not synthesize its own cholesterol, the worm ortholog, SBP-1 controls the expression of fatty acid Δ9 desaturase, elongase and synthase genes in *C. elegans* (McKay et al., 2003). Animals that do not have a functional *sbp-1* gene have decreased amount of body fat, exhibit growth defects and are sterile (Yang et al., 2006). These phenotypes can partially be rescued by adding specific fatty acids to the nematode growth medium; indicating that observed defects are direct consequences of impaired metabolic gene expression (Yang et al., 2006).

Another *C. elegans* metabolic TF is NHR-49, which belongs to the nuclear hormone receptor (NHR) family of TFs. NHR-49 is required for the expression of several fasting response genes, including *acs-2*, *acs-11*, and *hacd-1* (mitochondrial β-oxidation), *fat-7* (fatty acid desaturation), and *lbp-8* (lipid binding) (Van Gilst et al., 2005a). *nhr-49* mutants store abnormally high body fat

due to the decreased expression of β-oxidation genes and are short-lived because of the reduced expression of Δ9 desaturases (Van Gilst et al., 2005a).

Transcriptional coactivator complexes can modulate the regulatory output of TFs. Using a directed approach that aimed to find the components of the regulatory complexes with which NHR-49 engages, Taubert *et al.* found that MDT-15, the *C. elegans* ortholog of the human ARC105/MED15 subunit of the Mediator complex, interacts with NHR-49, potentially through its conserved KIX domain, and is required for the expression of several NHR-49 targets (Taubert et al., 2006). Contemporaneously, Yang *et al.* found that MDT-15 also interacts with SBP-1 and in the absence of MDT-15, the expression of SBP-1 targets is severely impaired (Taubert et al., 2006; Yang et al., 2006). *mdt-15* mutants also display noticeable growth defects, are sterile and short-lived.

Although NHR-49 and MDT-15 function together and are required for the expression of several fasting response genes, not all MDT-15 targets are regulated by NHR-49; indicating that MDT-15 likely interacts with additional TFs (Taubert et al., 2006).

NHRs are global regulators of metabolism

NHRs are a family of ligand-gated TFs. The canonical structure consists of two conserved domains; an N-terminal zinc finger DNA binding domain (DBD) and a C-terminal ligand binding domain (LBD) (Rastinejad et al., 1995; Huang et al., 2010). NHRs bind to DNA sequences called hormone response elements

(HREs) and can do so as monomers, homodimers or heterodimers (Sonoda et al., 2008). HREs consist of the typical sequence RGGTCA where R is a purine. NHR ligands are usually small hydrophobic molecules, such as steroids, fatty acids and xenobiotics (Chawla et al., 2001). Ligand binding can modulate the transcriptional output by switching the NHR from a transcriptionally inactive state to an active state or *vice versa.* Mechanistically, ligand binding causes a conformational change at the 12^{th} helix of the LBD, which essentially alters the receptor's ability to contact transcriptional coactivator or repressor complexes (Nagy and Schwabe, 2004; Sonoda et al., 2008). The evolution of the NHR superfamily, which can specifically recognize and bind small molecules and directly alter gene expression, is thought to have allowed metazoans to rapidly integrate signals from the metabolic state of the body as well as signals from the environment into gene expression. Having the ability to act as molecular switches, NHRs regulate almost all aspects of animal physiology; including carbohydrate and lipid metabolism, toxin metabolism, development and reproduction (Sonoda et al., 2008).

Based on phylogenetic analysis, NHRs are grouped into six subfamilies; NR1-NR6 (Nuclear Receptor Committee, 1999). Steroid receptors, such as the estrogen receptor (ER), and the androgen receptor belong to the NR3 subfamily, and are capable of binding their ligands with high-affinity (Chawla et al., 2001; Nagy and Schwabe, 2004). NR3 subfamily exists in vertebrate genomes, and has one member in the *Drosophila* genome; however it is absent in *C. elegans*.

Another notable absence in the *C. elegans* genome is the NR2B family, which includes retinoid X receptors (RXR) that bind to a variety of ligands (dietary lipids, xenobiotics) with low affinity (Chawla et al., 2001). RXRs are obligate heterodimers and partner with several other nuclear receptors from other subfamilies such as the peroxisome proliferator-activated receptors (PPARs), liver X receptor (LXR) and the farnesoid X receptor (FXR), which all belong to the NR1H group (Van Gilst et al., 2002).

Despite the losses of these subfamilies, remarkably the *C. elegans* genome encodes a total of 284 NHRs compared to 21 in *Drosophila* and 48 in humans (Maglich et al., 2001). Interestingly however, only 15 of these *C. elegans* NHRs show significant sequence similarity to the conserved vertebrate homologs, while the remaining 269 of them have diverged significantly. Most *C. elegans* NHRs are homologs of hepatocyte nuclear factor 4 (HNF4). HNF4 was initially found in crude rat liver extracts and later purification and cloning efforts showed that HNF4 presents a new class of nuclear receptors (NR2A) (Sladek et al., 1990). *Hnf4* null mouse fails to complete gastrulation in the early embryo, which results in embryonic lethality (Watt et al., 2003). Tissue specific *Hnf4* null mice have impaired liver function, and lipid homeostasis (Hayhurst et al., 2001). There are two variants of HNF4 in humans and only one in *Drosophila* (Palanker et al., 2009). In humans, HNF4α mutations lead to an early onset diabetic disorder, MODY1 (maturity onset diabetes of the young) (Yamagata et al., 1996). In *Drosophila*, dHNF4 mutants are sensitive to starvation and store higher levels

of fat, suggesting that dHNF4 responds to nutrient availability (Palanker et al., 2009). Thus, HNF4 likely plays an important role in post-developmental, metabolic GRNs in humans and flies. So far, only few *C. elegans* NHRs have been characterized, and for most their physiological and molecular functions remain unknown. Further, the evolutionary advantages of NHR family expansion have remained elusive, and the organization and functionality of *C. elegans* NHRs in the context of GRNs remain completely uncharacterized.

Synopsis

Physical and/or regulatory interactions between TFs and their target genes are essential to establish body plans of multicellular organisms during development, and these interactions have been studied extensively in the context of GRNs. The precise control of differential gene expression is also of critical importance to maintain energy homeostasis, and many metabolic disorders such as obesity and diabetes coincide with substantial changes in gene expression. Much work has focused on the GRNs that control metazoan development; however, the design principles and organization of the GRNs that control systems physiology remain largely unexplored.

In the following chapters, we present the mapping of the first gene-centered GRN that relates to *C. elegans* metabolism and physiology. This network contains numerous interactions between metabolic gene promoters, TFs and a cofactor. We describe the local architecture of the metabolic GRN and how that relates to the lipid homeostasis of the nematode. We also investigate several mapped interactions *in vivo*, and show their biological relevance. Lastly, we discuss the significance and implications of our findings, and comment about future perspectives of *C. elegans* metabolic GRNs.

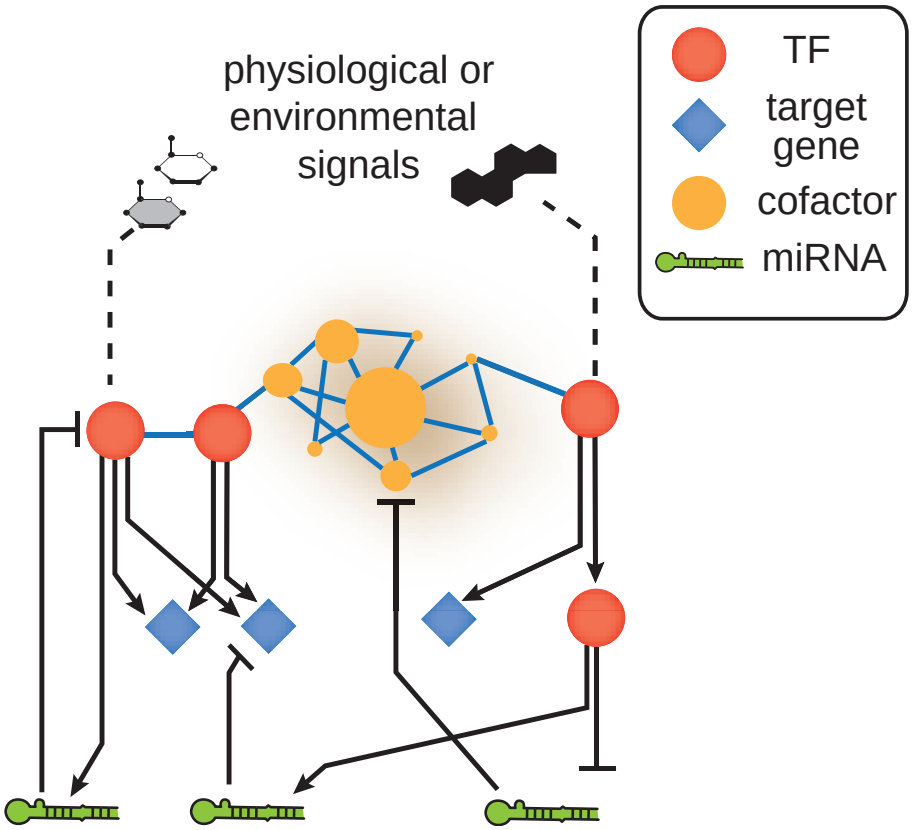

Figure 1.1 Cartoon illustrating hypothetical interactions in a GRN.

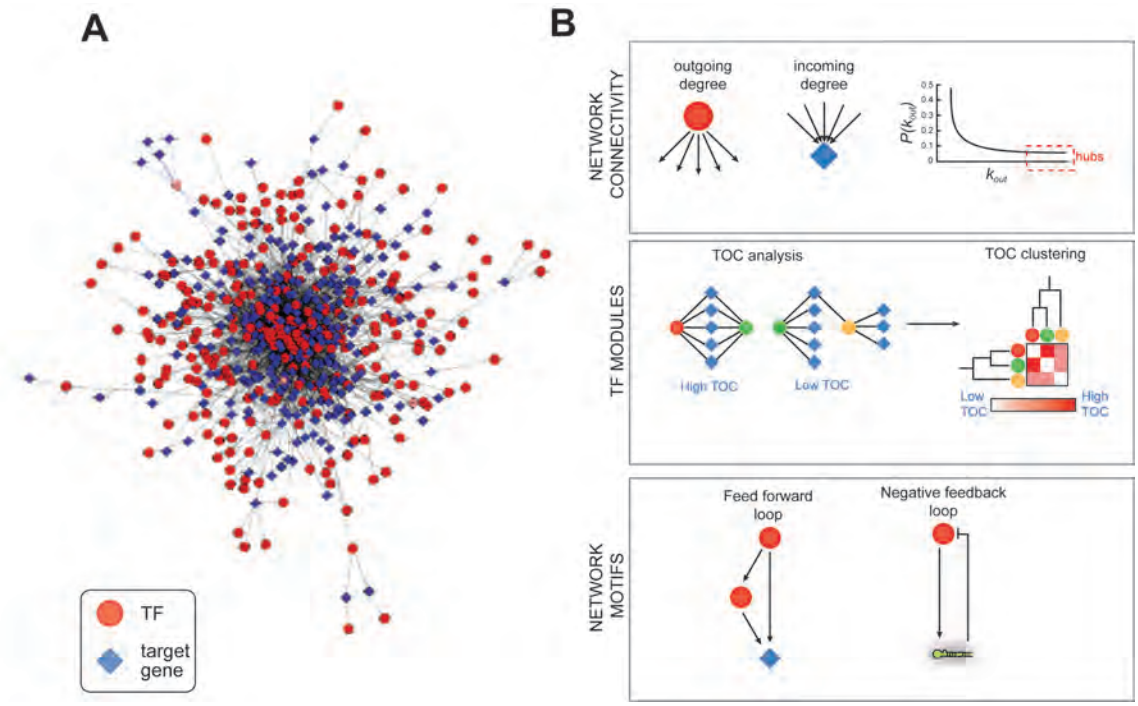

Figure 1.2 Features of GRNs.

(A) A compiled GRN of ~200 *C. elegans* gene promoters (Deplancke et al., 2006; Vermeirssen et al., 2007a; Arda et al., 2010). This model reflects the complexity of systems level gene regulation. The visualization was done using Cytoscape v.2.6 software with 'random layout' settings, which randomly distributes the nodes in a given network (Shannon et al., 2003).

(B) Analysis of incoming or outgoing degrees of individual nodes reveals network hubs *(top)*. TOC analysis and clustering can identify TF modules *(middle)*. Examples of network motifs that are overrepresented in GRNs *(bottom, hairpin structure – miRNA gene promoters).*

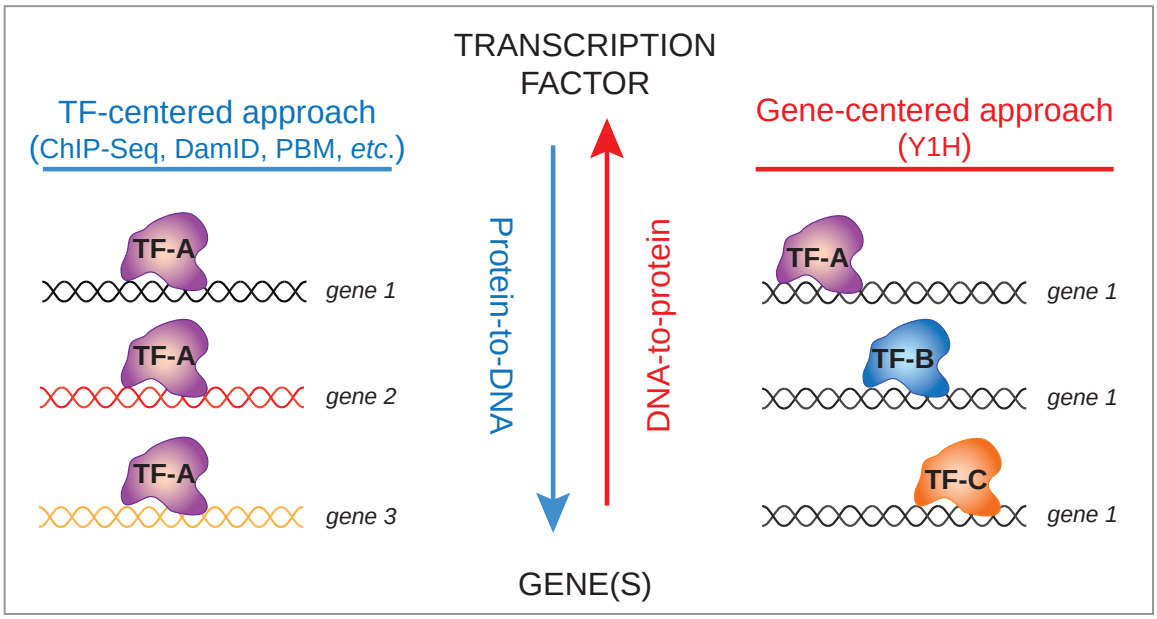

Figure 1.3 GRNs can be delineated using two conceptually different, but highly complementary approaches.

TF-centered methods start with a TF of interest and identify DNA fragments this TF binds to. Gene-centered methods start with a set of DNA fragments and identify the TFs these fragments interact with.

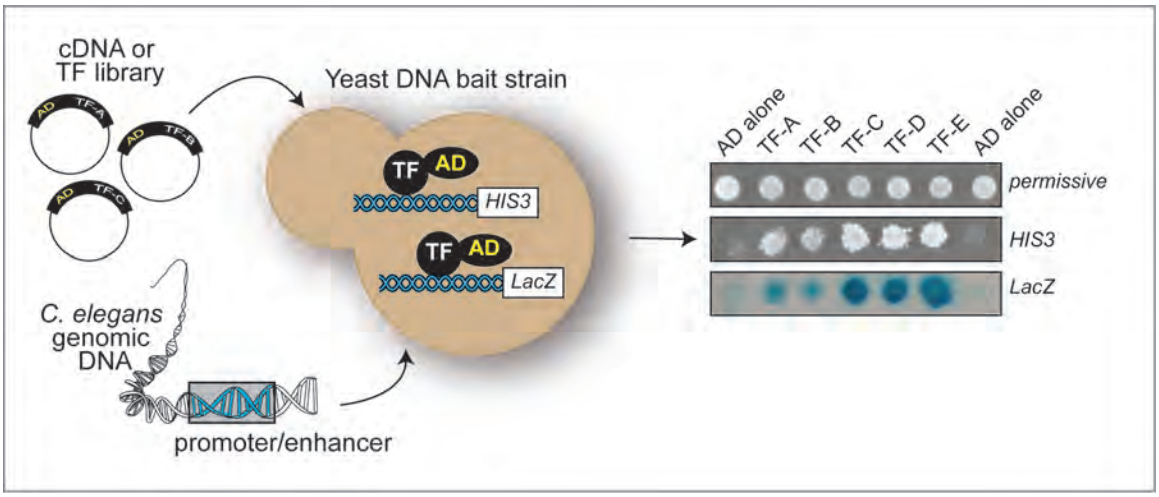

Figure 1.4 Gateway-compatible Y1H assays provide a convenient gene-centered method for GRN mapping.

A yeast DNA bait strain contains two reporter constructs integrated into its genome. Each construct contains the same DNA fragment, but different reporter genes, such as *HIS3* and *LacZ*. A cDNA or TF mini-library can be used as a prey resource, or collections of TFs can be tested individually. The panel on the right shows the readout of a Y1H experiment. AD alone indicates the negative control used to assess DNA bait background growth on media lacking histidine and containing 3-AT, and background bait coloring on a βGal assay. Different interacting TFs confer different interaction phenotypes; *i.e.* light blue *versus* dark blue.

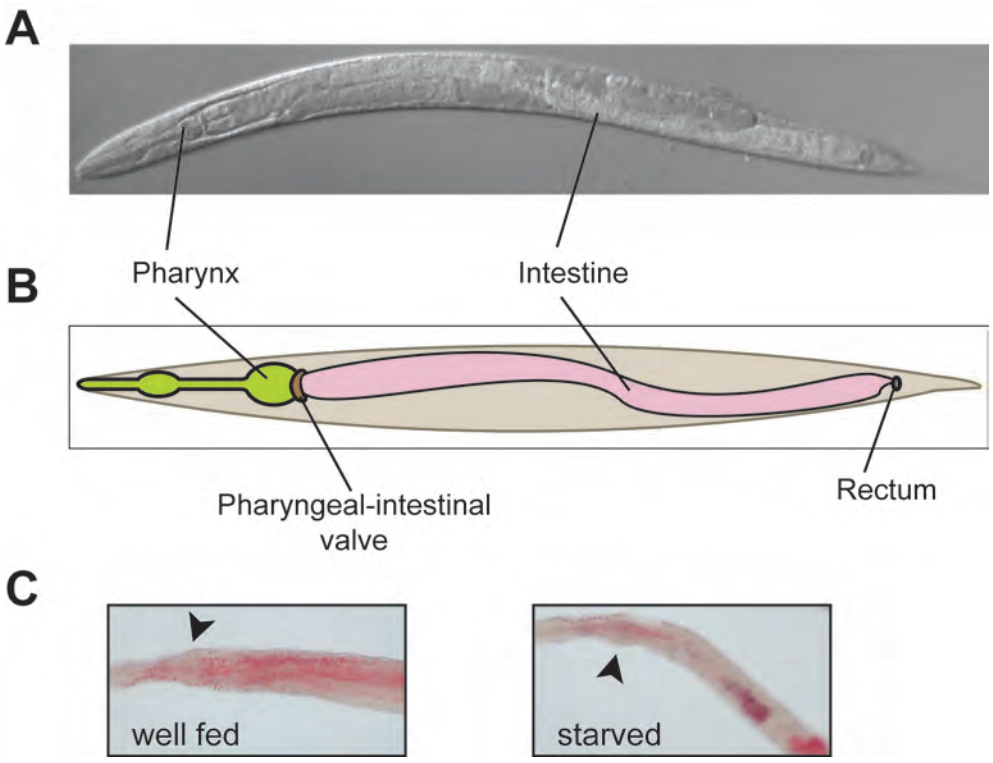

Figure 1.5 *C. elegans* as a model organism to study energy metabolism.

(A) Nomarski image of a *C. elegans* young adult hermaphrodite.

(B) Cartoon illustrating the alimentary system of *C. elegans.*

(C) Oil-Red-O lipid staining of a well fed or a starved nematode. Red droplets indicate stored body fat, which are visible in the well fed animal but not in the starved animal. Black arrowhead points to the posterior bulb of pharynx.

Preface To Chapter II

This chapter describes mapping of the first gene-centered regulatory network of *C. elegans* metabolic genes, analysis of its network organization, and functional role of nuclear hormone receptors that occur in this network in the maintenance of *C. elegans* lipid homeostasis. The work presented here was a collaboration with Drs. Stefan Taubert, Marc Van Gilst, and Keith Yamamoto, who provided the list of fasting response genes, and shared technical and scientific expertise. Ben Tsuda contributed to the optimization and generation of the Oil-Red-O lipid staining of data.

This chapter is a part of the following published manuscript, which was written by Dr. A.J. Marian Walhout and myself:

Arda, H. E., Taubert, S., MacNeil, L. T., Conine, C. C., Tsuda, B., Van Gilst, M., Sequerra, R., Doucette-Stamm, L., Yamamoto, K. R., and Walhout, A. J. M. (2010). Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. Mol. Syst. Biol *6*, 367.

CHAPTER II

Gene regulatory network of *C. elegans* metabolic genes

Abstract

Gene regulatory networks (GRNs) provide insights into the mechanisms of differential gene expression at a systems level. GRNs that relate to metazoan development have been studied extensively. However, little is still known about the design principles, organization and functionality of GRNs that control physiological processes such as metabolism, energy homeostasis and responses to environmental cues. Here, we report the first experimentally mapped metazoan GRN of *Caenorhabditis elegans* metabolic genes. This network is enriched for nuclear hormone receptors (NHRs). The NHR family has greatly expanded in nematodes; humans have 48 NHRs but *C. elegans* has 284, most of which are uncharacterized. We find that the *C. elegans* metabolic GRN is highly modular and that two GRN modules predominantly consist of NHRs. Network modularity has been proposed to facilitate a rapid response to different cues. Since NHRs are metabolic sensors that are poised to respond to ligands, this suggests that *C. elegans* GRNs evolved to enable rapid and adaptive responses to different cues by a concurrence of NHR family expansion and modular GRN wiring.

Introduction

Differential expression of metazoan genes in space and time is of critical importance to many biological processes, including the maintenance of energy balance. Differential gene expression is, at the first level, carried out by TFs that control the expression of their target genes by physically interacting with *cis*-regulatory DNA sequences, such as promoters and enhancers. Altogether, interactions between genes and their transcriptional regulators can be graphically represented in GRN models.

C. elegans is a powerful model organism to study metazoan GRNs. Several GRNs have been characterized to various degrees in *C. elegans*. These include protein-coding gene sets of endoderm, digestive tract, neurons, the C-lineage and the vulva, as well as microRNA and bHLH TF-encoding genes (Maduro and Rothman, 2002; Baugh et al., 2009; Deplancke et al., 2006a; Vermeirssen et al., 2007a; Martinez et al., 2008; Ririe et al., 2008; Grove et al., 2009). Despite these efforts however, little is known about the networks that control systems-level, post-developmental physiology in the nematode.

Investigation of the *C. elegans* energy pathways identified numerous genes that function in these pathways, and underscored the importance of transcriptional control in the regulation of *C. elegans* metabolism (Ashrafi et al., 2003; Van Gilst et al., 2005; Taubert et al., 2006). NHRs are important regulators of metabolism, and therefore are expected to control crucial aspects of

physiology in *C. elegans* as well. To date, few *C. elegans* NHRs have been characterized, despite the fact that *C. elegans* genome has 284 NHRs. In this chapter, we describe the mapping and the network analysis of the first gene-centered GRN that relates to *C. elegans* metabolism and physiology. We show that NHRs indeed are essential for the *C. elegans* energy homeostasis and they organize into functional TF modules.

Results

Compiling a set of metabolic genes

To gain insight into the organization and functionality of GRNs involved in systems physiology, we first selected a set of genes that have been implicated in *C. elegans* metabolism. Two thirds of the genes were identified in a genome-wide RNAi screen to find gene inactivations that caused an altered lipophilic dye staining pattern in both wild type animals and animals with multiple genetic backgrounds (Ashrafi et al., 2003). The lipophilic dye of choice was Nile Red. When used as a vital dye as opposed to fixative methods, Nile Red stains 'fat-containing lysosome-like organelles' in the *C. elegans* intestine (Schroeder et al., 2007; Rabbitts et al., 2008). Thus, the genes uncovered in the RNAi study may be involved in lipid metabolism, and/or in other types of metabolism such as the general catabolism of biomolecules. The other third of our gene set was identified in an effort to find metabolic genes whose expression is affected by food availability. These 'fasting response genes' give a robust transcriptional response upon short-term food withdrawal, and the regulation of some, but not all of these, is dependent on the nuclear receptor, NHR-49 (Van Gilst et al., 2005). Hereafter we collectively refer to these genes as 'metabolic genes' (Figure 2.1).

Mapping the gene-centered metabolic GRN

To identify TFs that can interact with metabolic genes, we cloned the promoters of 71 metabolic genes upstream of the Y1H reporter genes *HIS3* and *lacZ,* and integrated the resulting *promoter::reporter* constructs into the yeast genome to create Y1H 'bait' strains (Deplancke et al., 2004, 2006b) (Table 2.1). We screened each bait strain *versus* a cDNA library (Walhout et al., 2000b), and a TF mini-library (Deplancke et al., 2006a). Subsequently, we tested each bait strain *versus* each TF identified by direct transformation, both to confirm interactions and to identify additional ones that were missed in the library screens (Table 2.2). We then scored and filtered the Y1H interactions as described (Vermeirssen et al., 2007a) to minimize the inclusion of false positives. Finally, we used Cytoscape (Shannon et al., 2003) to combine all interactions into a GRN graph model (Table 2.3 and Figure 2.2A).

In total, the metabolic GRN contains 508 interactions between 69 metabolic gene promoters and 100 TFs (Figure 2.2A).

Network analysis of the metabolic GRN

All components in the metabolic GRN are connected in a single graph due to the presence of both highly connected promoters and highly connected TFs. The network is densely wired and the global structure is similar to that of other gene-centered GRNs (data not shown) (Deplancke et al., 2006a; Vermeirssen et al., 2007a; Martinez et al., 2008). However, we did observe a striking difference:

more than a quarter of the TFs retrieved here are NHRs, which is significantly more than in the digestive tract, neuronal and microRNA networks (Deplancke et al., 2006a; Vermeirssen et al., 2007a; Martinez et al., 2008) (Figure 2.2B).

This is exciting because, as mentioned above, NHRs can function as metabolic sensors. The retrieval of many NHRs suggests that the expansion of this family relates to metabolic functionality. The difference with the digestive tract network is relatively modest *(p=0.05)*, which is probably because the intestine is the most important metabolic tissue. Overall, 41 of the 69 promoters (~60%) interacted with one or more NHR, suggesting that the promoters of *C. elegans* metabolic genes may have an inherent preference for NHRs, and that multiple NHRs may regulate metabolic gene expression.

NHRs organize into functional modules

Systems-level GRNs can capture hundreds of interactions between numerous genes and their regulators, and such networks are often too complex for manual analysis. Instead, mathematical and computational methods can be used to investigate the design principles and organization of GRNs. These principles can then be related to biological functionality. For instance, GRNs can be decomposed into 'modules'; highly interconnected network neighborhoods consisting of nodes with similar functions (Ravasz et al., 2002; Vermeirssen et al., 2007a). Such modular network organization has been proposed to increase

the adaptability of a system and to allow rapid and robust informational flow through the network (Ravasz et al., 2002; Babu et al., 2004).

To examine whether the *C. elegans* metabolic GRN has a modular architecture we performed topological overlap coefficient (TOC) analysis (Figure 1.2) (Vermeirssen et al., 2007a). For each TF pair, we calculated a TOC score based on the number of target genes they share in the metabolic GRN, and clustered the TFs with similar TOC scores to identify TF modules. After TOC clustering, we found that the metabolic GRN is highly modular as it contains five TF modules (I-V) (Figure 2.3A).

This is more than we have observed previously; the neuronal network consisted of only two TF modules, whereas the microRNA network did not contain any, (Vermeirssen et al., 2007a; Martinez et al., 2008) data not shown). Interestingly, ~60% (16 of 27) of all NHRs in the network are located in either one of two modules (modules II and III), and each of these modules consists predominantly of NHRs (66% each, Table 2.4, Figure 2.3B and Figure 2.4).

Lipid phenotype analyses of the NHRs that occur in modules

The observation that NHRs are wired into GRN modules that share target genes leads to the prediction that either: 1) one or few of them are involved in metabolic regulation *in vivo*; 2) they all act redundantly, or 3) they all function in the regulation of systems physiology. The majority of the target genes of the NHRs that participate in modules have a metabolic phenotype as judged by an

increase or decrease in Nile Red staining (Ashrafi et al., 2003). Thus, we performed systematic Nile Red staining upon reduction of the activity of different NHRs by RNAi. Several NHRs in module II are essential for *C. elegans* development (Kamath et al., 2003) and could therefore not be examined. Nonetheless, RNAi of one NHR in module II and most NHRs in module III resulted in an increase in Nile Red staining (Figures 2.5A and 2.5B). To further analyze changes in fat depots, we performed Oil-Red-O staining of animals subjected to *nhr(RNAi)*. We found that RNAi of most NHRs resulted in increased Oil-Red-O staining, indicating that most of these NHRs indeed regulate fat storage or catabolism (Figures 2.6A and 2.6B). In *Drosophila* the single HNF4 homolog is responsible for the regulation of fat storage (Palanker et al., 2009). In contrast, our findings indicate that, in *C. elegans*, multiple HNF4-type NHRs share this function, even after duplication and divergence. Altogether these findings demonstrate that module III contains functionally related NHRs that all regulate *C. elegans* physiology. The fact that the NHRs in module III are dispensable for development suggests that these function in post-developmental physiology, for instance to respond to environmental or dietary cues, whereas the NHRs in module II may regulate metabolic gene expression during development.

Discussion

In this chapter, we present the first experimentally mapped, metabolic GRN in a metazoan model system. This network contains hundreds of protein-DNA interactions between several metabolic genes and several TFs. We find that the GRN is enriched for NHRs compared to other gene-centered networks and that it is highly modular. Two modules mainly contain NHRs, and remarkably, most of these NHRs confer a metabolic phenotype.

It is important to note that the metabolic GRN is not yet complete. First, we used only a subset of metabolic genes. Second, we have so far exclusively focused on gene promoters and it is likely that other cis-regulatory elements such as enhancers may be involved in differential metabolic gene expression as well. Third, not all interactions are detectable by Y1H assays; i.e. TFs for which we do not have a (correct) clone or that are underrepresented in the cDNA library will not be retrieved, and the system is not yet adapted to identify TF heterodimers. This may explain, at least in part, why we did not retrieve NHR-49, which dimerizes with more than 20 other NHRs (Vermeirssen et al., 2007b) (see also below).

In addition to missing interactions, the network may also contain interactions that are not biologically relevant. These may be false positives, or alternatively do occur in vivo, but do not lead to an observable biological consequence. Indeed, it has been observed that TFs can interact with genomic

locations without an apparent regulatory effect (Zeitlinger et al., 2007; Li et al., 2008). Further, it may be that interactions do have a biological effect but that this cannot be detected in vivo (i.e. it could be a false negative of the validation assay). For example, interactions that occur in a few cells, only under specific environmental conditions, or with small regulatory effects may be difficult to detect. Finally, it is of course also possible that some interactions that we retrieved really do not occur in vivo and are actual false positives. We have aimed to minimize the inclusion of false interactions in our network by ensuring the technical quality of our experiments; the DNA baits are integrated into the yeast genome and are, therefore, both present at fixed copy number and in a chromatinized state. This is essential to avoid spurious interactions (Deplancke et al., 2004). In addition, we scored and filtered all our Y1H data to minimize the inclusion of 'sticky' TFs or interactions that were retrieved with highly autoactive promoters (Vermeirssen et al., 2007a). Despite some of these potential limitations, the results presented here, and in our other studies, demonstrate the use of Y1H assays to identify meaningful TF-target gene interactions (Deplancke et al., 2004, 2006a; Vermeirssen et al., 2007a; Martinez et al., 2008; Reece-Hoyes et al., 2009). Most importantly, the Y1H system uniquely enabled us to delineate a GRN pertaining to physiology because it can be used with multiple genes to retrieve multiple TFs. Indeed, more than 10% of all predicted C. elegans TFs (Reece-Hoyes et al., 2005) were retrieved in our network.

NHRs are known regulators of metabolic gene expression and this enrichment, therefore, provides global support for the overall quality of our network. We found that NHRs organize into two TF modules in the network. Network modularity has been proposed to facilitate rapid and robust changes in gene expression (Ravasz et al., 2002; Babu et al., 2004). Previously, it has been shown that biochemical networks composed of reactions between metabolic enzymes and their substrates are modular as well (Ravasz et al., 2002). Thus, we propose that *C. elegans* has acquired modularity in multiple layers of its physiological networks.

Materials and Methods

Promoter cloning

Gene promoter was defined as the intergenic region upstream of the translational start site with a minimum length of 300 bp and a maximum length of 2 kb. The promoters were cloned as described (Dupuy et al., 2004). Briefly, promoters were amplified from *C. elegans* (N2) genomic DNA by PCR, cloned into two Y1H reporter Destination vectors via Gateway cloning and integrated into the genome of *S. cerevisiae* (YM4271) to create Y1H bait strains as described (Deplancke et al., 2006b). In total, 71 Y1H bait strains were successfully generated and screened in Y1H assays. Primer sequences and detailed information about Y1H bait strains are provided in Table 2.1.

Y1H assays

Y1H assays were performed as described (Deplancke et al., 2006b; Vermeirssen et al., 2007a; Martinez et al., 2008). Two different prey libraries AD-wrmcDNA (Walhout et al., 2000b), and AD-TF mini library (Deplancke et al., 2006a) were used to screen each Y1H bait strain to identify interacting TFs. All interactions were retested in fresh yeast by PCR/gap repair. The ORFs of preys that retested were PCR amplified from corresponding yeast strains and were sequenced by Agencourt Bioscience Corporation. Interacting TFs were identified using the BLASTX algorithm. In total, 306 unique Interaction Sequence Tags

(ISTs) were obtained (Walhout et al., 2000a). Y1H matrix experiments were performed by transforming available prey clones of all interacting TFs obtained in the screens, and several TFs found previously (Deplancke et al., 2006a; Vermeirssen et al., 2007a; Martinez et al., 2008) into each Y1H bait strain individually (174 preys were used in total, Table 2.2). Each fat gene promoter-TF interaction was evaluated using a standardized and stringent Y1H scoring and filtering system (Vermeirssen et al., 2007a). Only interactions with a score ≥ 5 were considered (Table 2.3). All interactions are available in the EDGEdb database (Barrasa et al., 2007).

Topological overlap coefficient analysis

TOC analysis and clustering was performed as described (Vermeirssen et al., 2007a).

RNAi experiments

We generated RNAi constructs for *nhr-86, nhr-234, nhr-109, nhr-273, nhr-41, nhr-178, nhr-79, nhr-12, nhr-28, nhr-70,* and *nhr-102* by transferring the full-length ORFs from ORFeome v3.1 into pL4440-Dest-RNAi (Rual et al., 2004) via Gateway cloning. Additional RNAi clones were cherry-picked from the C. elegans RNAi library (Kamath et al., 2003). HT115(DE3) bacterial strains carrying RNAi constructs were grown in Luria Broth containing ampicillin (50 µg/mL) at 37°C for 8-10 hours until they reach OD600 ~1.0. Bacteria were pelleted by centrifugation,

washed once with M9 buffer, resuspended in M9 buffer, and seeded on NGM plates containing 5 mM IPTG and 50 µg/mL ampicillin. Bacteria were induced overnight at 22°C. The next day ~30 synchronized L1 larvae were placed on the plates. All RNAi clones were verified by sequencing.

Microscopy

Nomarski and fluorescence images were obtained using a Zeiss Axioscope 2+ microscope. Nile Red fluorescence images were taken using a rhodamine filter (excitation 525-555 nm, emission 575-630 nm). Animals were placed into a drop of 0.1% sodium azide in M9 buffer on a fresh 2% agarose pad for observation. Oil-Red-O images were acquired using Zeiss AxioCam HRc color CCD camera.

Vital Nile Red staining and quantification

Nile Red staining was performed as described (Ashrafi et al., 2003). Briefly, Nile Red powder (N-1142 Molecular Probes, Invitrogen) was dissolved in acetone to make a 0.5 mg/mL stock solution, which was kept at -20°C. This stock solution was then diluted 1:500 in 1X phosphate buffered saline (PBS), and overlaid on top of NGM plates either seeded with OP50 or with RNAi bacteria (HT115). After the plates were equilibrated overnight at 22°C, synchronized L1 stage animals were placed onto these plates. When animals reached the L4 stage (approximately 36 hours later at 20°C), the phenotype assessment was

done by fluorescence microscopy (see below for details). To quantify Nile Red pixel intensities, we followed the procedure outlined in (Srinivasan et al., 2008). Identities of images were masked while recording data to prevent observer bias. Statistical analysis of RNAi coupled Nile Red experiments of multiple NHRs was done by one-way ANOVA, followed by Dunnett's Multiple Comparison post-test, using GraphPad Prism v.5.00 software.

Oil-Red-O staining and quantification

Oil-Red-O staining was performed as described in (Soukas et al., 2009). *Quantification of Oil-Red-O staining by image processing:* RGB images of stage-synchronized, 1-day old adult animals (n=3) were acquired using identical bright field settings to keep background illumination constant across different samples. These images were digitally oriented and cropped to include only the anterior part of each worm, standardizing the area to be processed. The RGB composite images were then split into each contributing color channel; red, green and blue. Of the three, the green channel was selected to further analyze pixel intensities, since it gives the best contrast for the red Oil-Red-O stain. The images were inverted to make the background dark, and thresholded to separate stained areas from the background using ImageJ software. The number of pixels corresponding to various pixel intensities from the thresholded areas were counted and plotted.

Acknowledgements

This work was supported by NIH grants DK068429 and DK071713, and by Grant number 5 P30 DK32520 from the National Institute of Diabetes and Digestive and Kidney Diseases to AJMW and NIH grant CA020535 to KRY. We thank members of our laboratories, Heidi Tissenbaum, YongXu Wang, Eric Baehrecke, Pedro Batista and Job Dekker for comments and discussions, Inma Barrasa for statistical advice, and Eunsoo Kwon for the *daf-2 RNAi* clone.

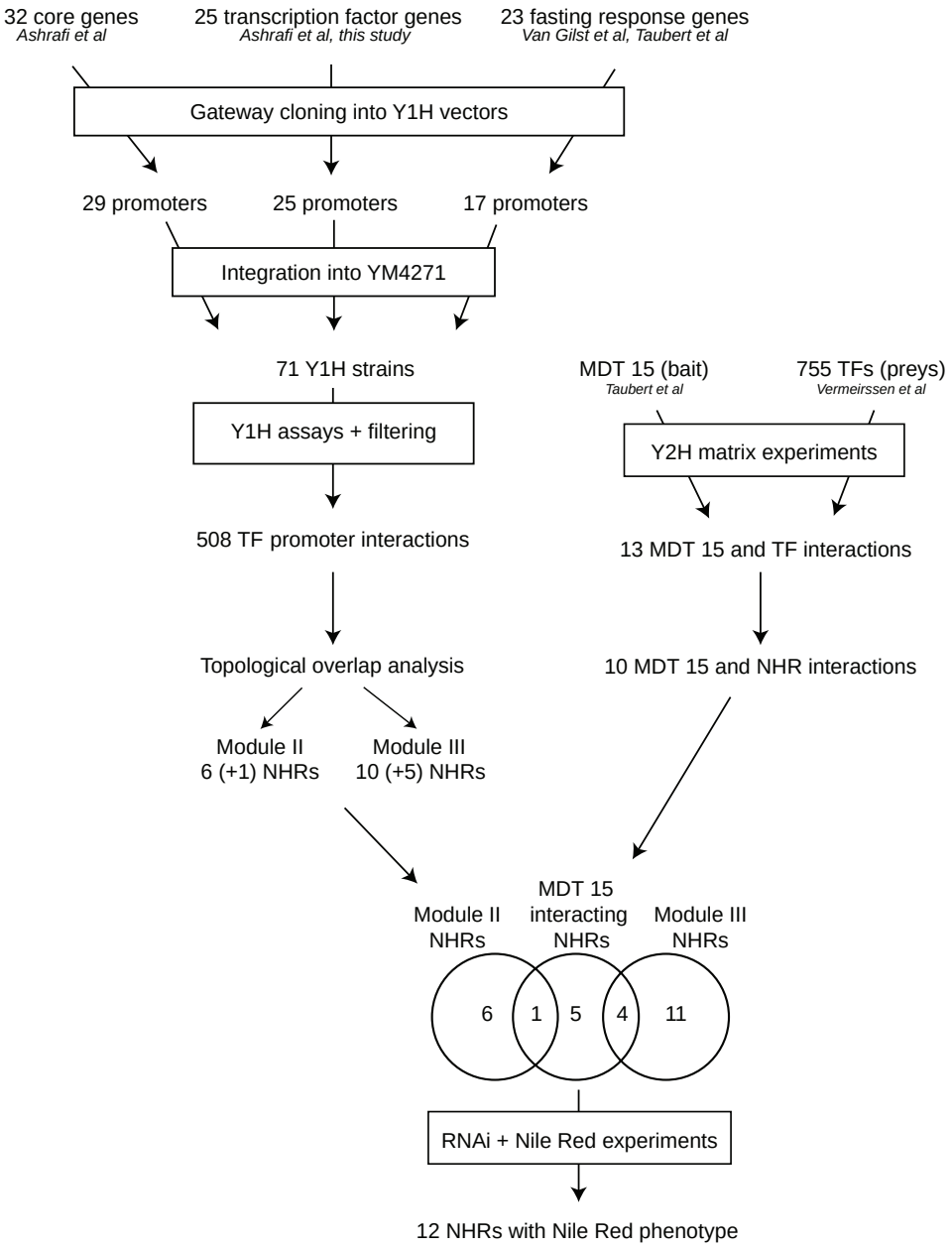

Figure 2.1 Schematic overview of this study.

Numbers in parentheses indicate NHRs that specifically bind to a single target that is part of the module.

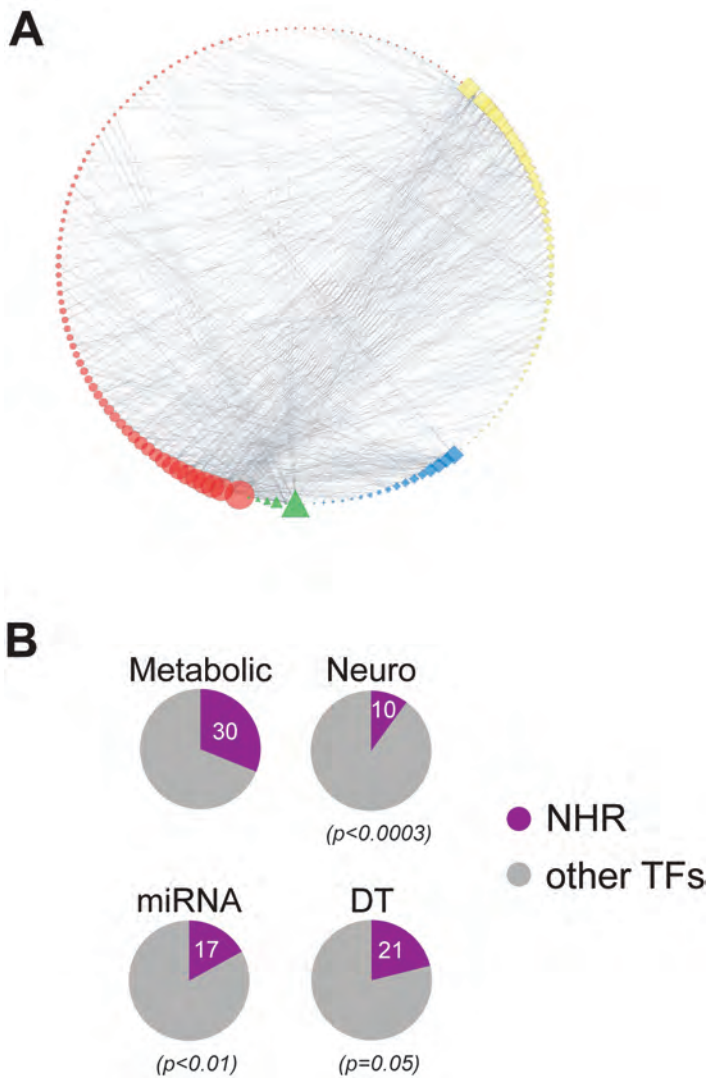

Figure 2.2 A *C. elegans* metabolic GRN.

(A) Protein-DNA interaction network of *C. elegans* metabolic genes. Red circles - TFs, yellow diamonds - promoters of genes with Nile Red phenotypes, blue diamonds - promoters of fasting response genes, green triangles - TFs whose promoters were also analyzed, grey lines - protein-DNA interactions identified by Y1H assays.

(B) Pie charts representing the percentage of NHRs in different gene-centered GRNs. Neuro – neuronal (Vermeirssen et al., 2007a), miRNA – microRNAs (Martinez et al., 2008), DT - digestive tract (Deplancke et al., 2006a). Putative novel TFs (that were specifically retrieved but that do not possess a known DNA binding domain (Deplancke et al., 2006a) were omitted.

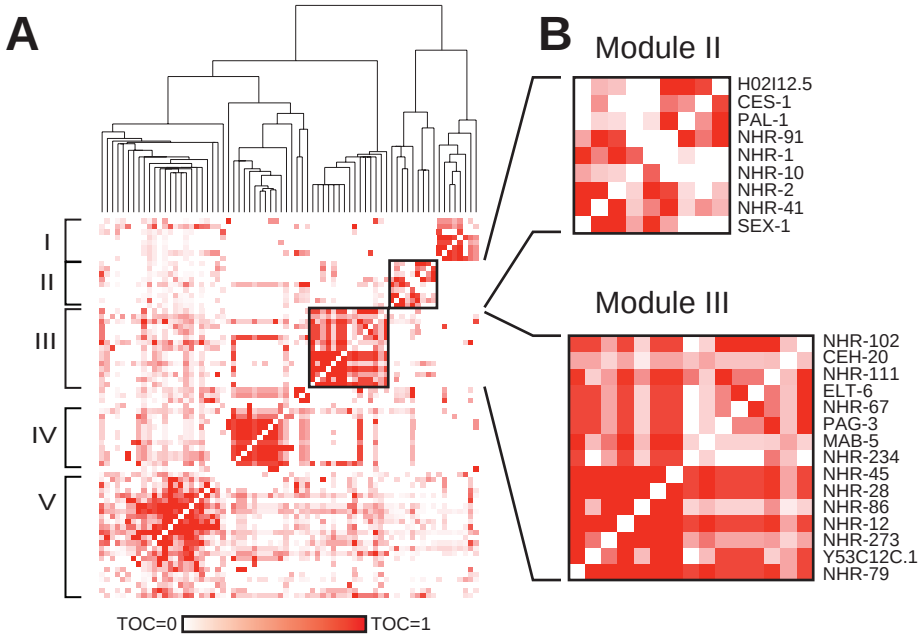

Figure 2.3 Modularity analysis of the metabolic GRN.

(A) TOC clustering matrix of TFs in the metabolic gene network identifies GRN modules (Roman numbers). The matrix is symmetrical across the white diagonal. Each cell that is at the intersection of two TFs represents the calculated TOC score for that TF pair. TOC scores are color coded as shown at the bottom of the panel.

(B) Enlarged views of Modules II and III, which predominantly contain NHRs.

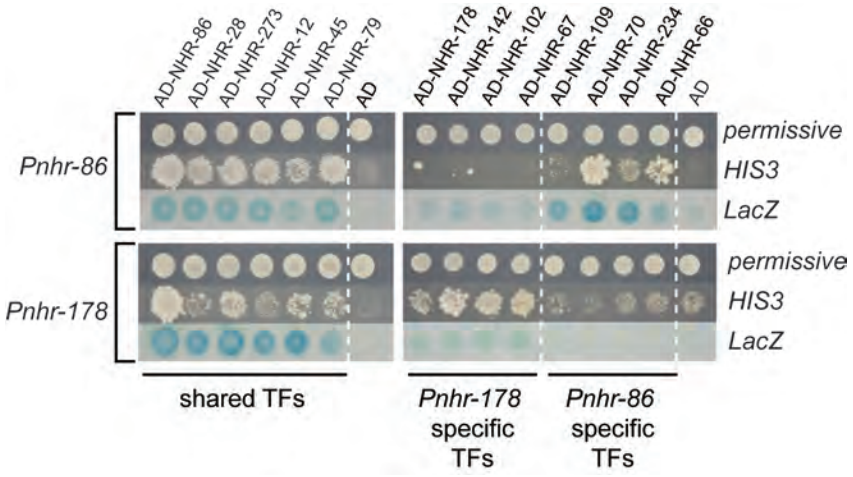

Figure 2.4 A Y1H experiment using NHRs (preys) and target NHR promoters (baits) that occur in Module III.

Each yeast spot contains indicated Gal4AD-NHR plasmid transformed into Y1H baits *Pnhr-86* (top), and *Pnhr-178* (bottom). The spots labeled as AD have no inserts, hence are negative controls. All spots grow on permissive plates (permissive); however only the AD-NHRs that bind either *Pnhr-86* or *Pnhr-178* enable growth on selective plates (*HIS3*) and activate *lacZ* reporter (*lacZ*). Y1H baits *Pnhr-86* and *Pnhr-178* have common (left) as well as specific interacting TFs (right).

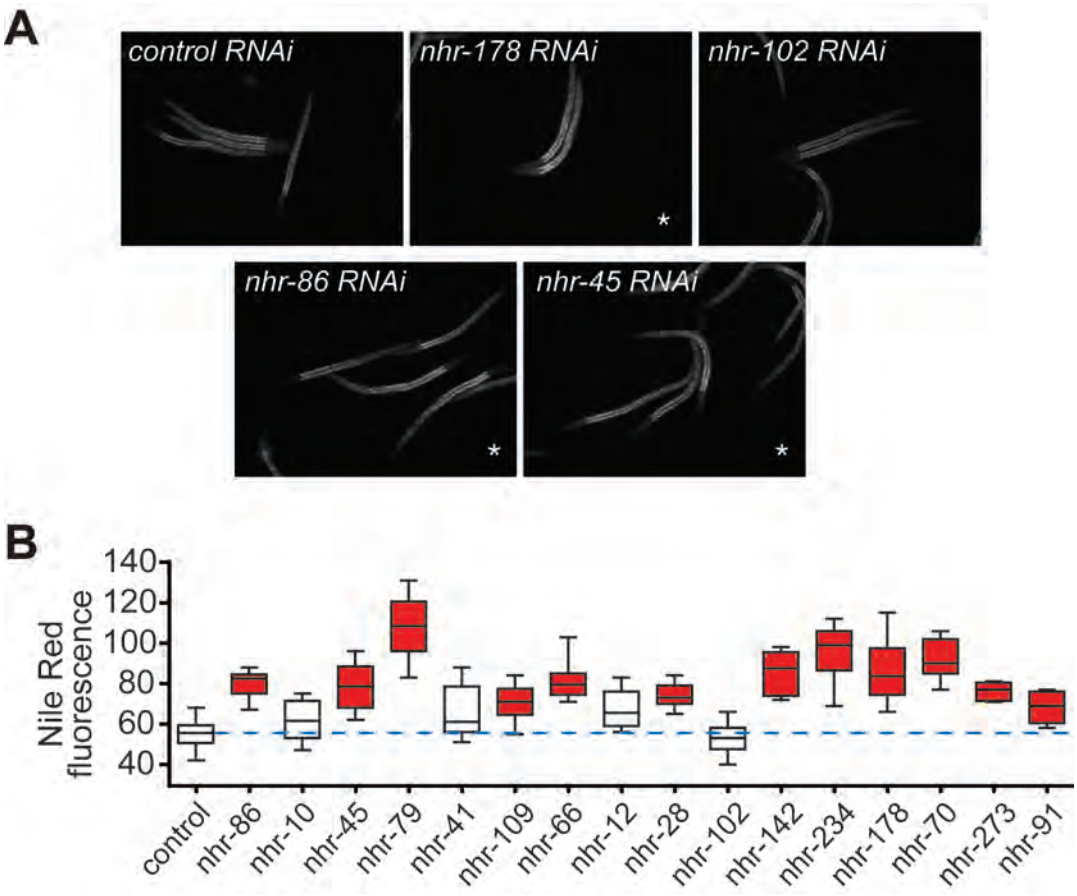

Figure 2.5 Analysis of vital Nile Red staining of NHR knockdowns.

(A) Examples of Nile Red phenotypes observed upon NHR inactivation by RNAi. Asterisks indicate significant changes in Nile Red fluorescence compared to control animals (notice the increased fluorescent intensity).

(B) Quantification of Nile Red staining coupled RNAi experiments. Red box plots - statistically significant changes in Nile Red staining, white box plots - no significant change, dashed line - median Nile Red fluorescence in control RNAi animals. In each box plot, the central bar indicates the median, the edges of the box indicate the 25th and 75th percentiles, and the whiskers extend to the most extreme data points. For details of the statistical analysis, see Materials and Methods.

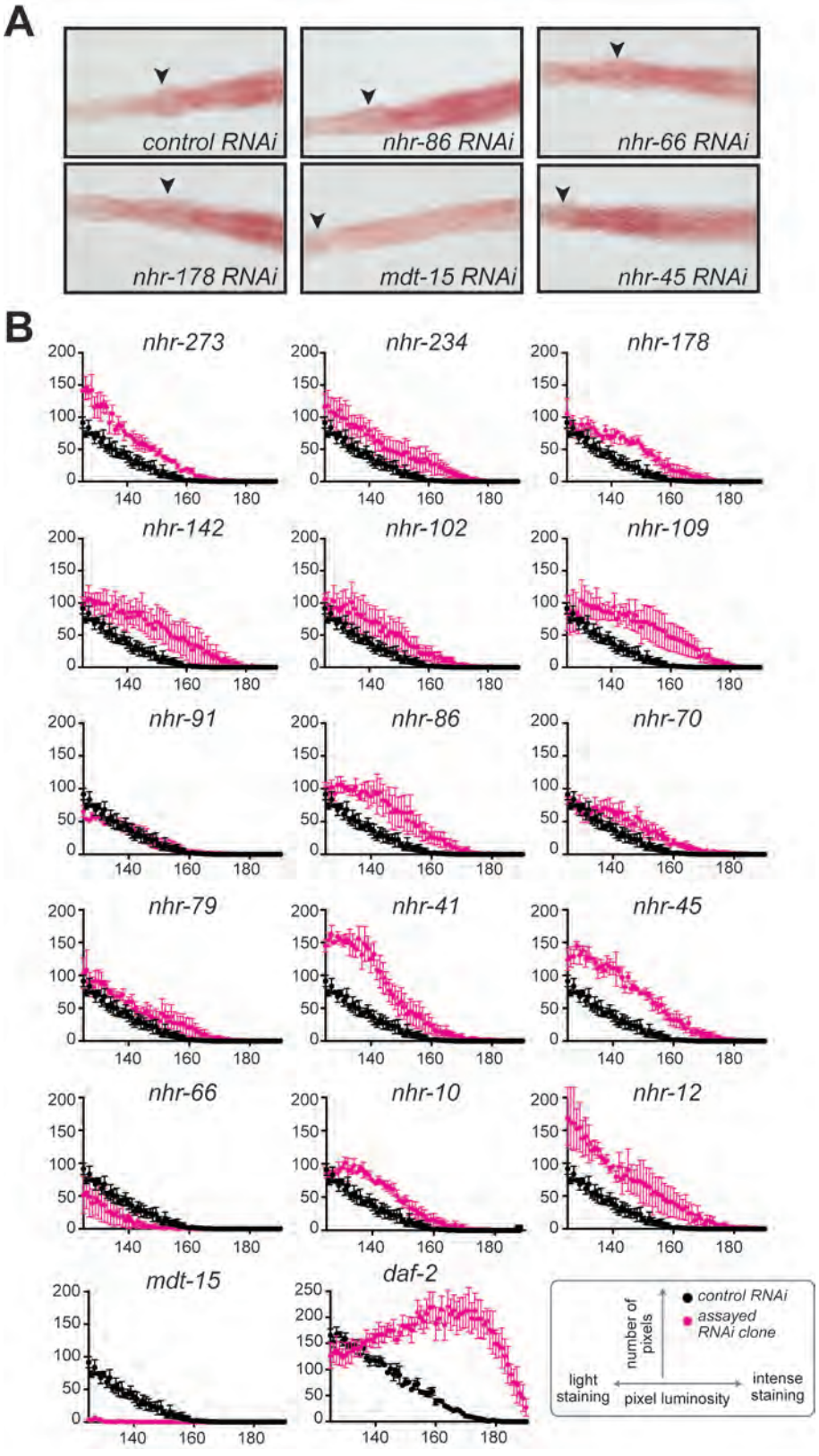
A
control RNAi
nhr-86 RNAi
nhr-66 RNAi
nhr-178 RNAi
mdt-15 RNAi
nhr-45 RNAi
B
nhr-273
nhr-234
nhr-178
nhr-142
nhr-102
nhr-109
nhr-91
nhr-86
nhr-70
nhr-79
nhr-41
nhr-45
nhr-66
nhr-10
nhr-12
mdt-15
daf-2
control RNAi
assayed RNAi clone
number of pixels
light staining
pixel luminosity
intense staining

Figure 2.6 Oil-Red-O staining analysis of NHR knockdown experiments.

(A) Examples of Oil-Red-O stained animals. Anterior part of each animal is to the left. Black arrowheads point to the posterior bulb of the pharynx.

(B) Quantifications of Oil-Red-O staining experiments. The inset at the bottom right of the figure describes graph axes. Error bars indicate the standard error of the mean (SEM). The two graphs at the bottom are negative (*mdt-15* RNAi) and positive controls (*daf-2* RNAi).

Table 2.1 Information on the metabolic gene promoters analyzed in this study

Promoter name	*WB Gene ID*	*Public name*	*Sequence name*	[1] *GW-Forward Primer*	[2] *GW-Reverse Primer*	*Comments*	*Promoter cloned?*	*Y1H interactions after filtering?*
Pacs-11	WBGene00018269	*acs-11*	F41C3.3	GGGGACAACTTTGTATAGAAAAGTTGAAAAGTTCAGGATCTTCCCA	GGGGACTGCTTTTTTGTACAAACTTGTCATTTCTGAAATATCTAATTTTCTGT		yes	yes
Pcyp-35A3	WBGene00019565	*cyp-35A3*	K09D9.2	GGGGACAACTTTGTATAGAAAAGTTGAAAGTTTACTTTTATATTTGGTGCTA	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTGAGAATTAAAAATTGAAAATA		yes	yes
Pech-6	WBGene00001155	*ech-6*	T05G5.6	GGGGACAACTTTGTATAGAAAAGTTGAAATTTTTTGTAGTGGCCTATGAAAGT	GGGGACTGCTTTTTTGTACAAACTTGTCATCCTGAAAATCAAAAAAAAATTAAATAA		yes	yes
Pnhr-8	WBGene00003607	*nhr-8*	F33D4.1	GGGGACAACTTTGTATAGAAAAGTTGAACTTTCATCTTTGTTATAGAACCGA	GGGGACTGCTTTTTTGTACAAACTTGTCATGGAATGACGAAATTTTTGTTTTAG		yes	yes
Pgei-7	WBGene00001564	*gei-7*	C05E4.9	GGGGACAACTTTGTATAGAAAAGTTGAAGGTATAAATTTTACAACAAAAAAATCGATG	GGGGACTGCTTTTTTGTACAAACTTGTCATTGTGTGTGTCTTTGTATAATAAGC		yes	yes
PF59F5.2	WBGene00010341	*F59F5.2*	F59F5.2	GGGGACAACTTTGTATAGAAAAGTTGAAGTACATATAATTCACTTTGCC	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTGAACGAGTTTACCTG		yes	yes
Pnhr-88	WBGene00003678	*nhr-88*	K08A2.5	GGGGACAACTTTGTATAGAAAAGTTGAATGATTAAACGTTATTTTGGAAAAAT	GGGGACTGCTTTTTTGTACAAACTTGTCATGATGCTGCAAAAAAAAGAAGATG		yes	yes
Pepn-1	WBGene00001329	*epn-1*	T04C10.2	GGGGACAACTTTGTATAGAAAAGTTGAGCACACAGAGTCTAATAGCTAA	GGGGACTGCTTTTTTGTACAAACTTGTCATGTTTGCTCCCCCGCCT		yes	yes
PZK593.3	WBGene00014003	*ZK593.3*	ZK593.3	GGGGACAACTTTGTATAGAAAAGTTGAGGCAGATCAATCAATTTTTAGGT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTGATTAACGAATGACT		yes	yes
Pmdt-15	WBGene00007016	*mdt-15*	R12B2.5	GGGGACAACTTTGTATAGAAAAGTTGATACTAACCACCACATATCTTCC	GGGGACTGCTTTTTTGTACAAACTTGTCATCTTTCTACTCTGCTCTTGTTTTT		yes	yes
PC06E7.3	WBGene00015540	*C06E7.3*	C06E7.3	GGGGACAACTTTGTATAGAAAAGTTGATATATGGCTCGGAACTTT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTATTTCGTTGATGGATGTTT		yes	yes
Pacdh-1	WBGene00016943	*acdh-1*	C55B7.4	GGGGACAACTTTGTATAGAAAAGTTGATCTTTACTTTCTAATTTCTATATAATTT	GGGGACTGCTTTTTTGTACAAACTTGTCATTGTGAAGGTGAAAGTAATTAGTG		yes	yes

[1] *att*B4tail: GGGGACAACTTTGTATAGAAAAGTTG
[2] *att*B1R tail: GGGGACTGCTTTTTTGTACAAACTTGTCAT

Promoter name	*WB Gene ID*	*Public name*	*Sequence name*	[1] *GW-Forward Primer*	[2] *GW-Reverse Primer*	*Comments*	*Promoter cloned?*	*Y1H interactions after filtering?*
Pnhr-49	WBGene00003639	*nhr-49*	K10C3.6	GGGGACAACTTTGTATAGAAAAGTTGATCTTTCTCTACCTCTTTTCATCT	GGGGACTGCTTTTTTGTACAAACTTGTCATGACTGAAACAATTAAACTTTTTACAAC		yes	yes
PT09F3.1	WBGene00011661	*T09F3.1*	T09F3.1	GGGGACAACTTTGTATAGAAAAGTTGATCTTTTGGCAAGTCTAGAAATCTGT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTTGAGTGTGTGTCAGGT		yes	yes
Psrj-34	WBGene00005619	*srj-34*	T07C12.5	GGGGACAACTTTGTATAGAAAAGTTGATGTCTTGGGTTGAATAATG	GGGGACTGCTTTTTTGTACAAACTTGTCATGTGAGCTGACCCTAGAGATTT		yes	yes
PF55B11.4	WBGene00010086	*F55B11.4*	F55B11.4	GGGGACAACTTTGTATAGAAAAGTTGATTATACTGATTGTTTTCTATATTTG	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTTCTGAAAATATAAACATTTTGA		yes	yes
Pelo-2	WBGene00001240	*elo-2*	F11E6.5	GGGGACAACTTTGTATAGAAAAGTTGATTGGTGAGACCAAATGTGTAT	GGGGACTGCTTTTTTGTACAAACTTGTCATTGTGATTACCTGCAAATTTCCACGA		yes	yes
Plbp-8	WBGene00002260	*lbp-8*	T22G5.6	GGGGACAACTTTGTATAGAAAAGTTGATTTTCATATTCAATAAACCTTGT	GGGGACTGCTTTTTTGTACAAACTTGTCATAAGGAAGGACTGAATTAATTGAG		yes	yes
PC09G9.7	WBGene00007496	*C09G9.7*	C09G9.7	GGGGACAACTTTGTATAGAAAAGTTGCAATATCTGCGGCCGTAAACTTT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTGACAGAATTATGATGATAAA		yes	yes
PF13D11.1	WBGene00017427	*F13D11.1*	F13D11.1	GGGGACAACTTTGTATAGAAAAGTTGCATCTAAACTACTTGTCCGCT	GGGGACTGCTTTTTTGTACAAACTTGTCATTATAAGGAGTTTTTGGAG		yes	yes
Pfat-6	WBGene00001398	*fat-6*	VZK822L.1	GGGGACAACTTTGTATAGAAAAGTTGCCGGAAGTTGCAGAAAGT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTACTGTCTGTTTTCTTTCTGAAAATTTTG		yes	yes
Pnhr-25	WBGene00003623	*nhr-25*	F11C1.6	GGGGACAACTTTGTATAGAAAAGTTGCCGGTTGTTGCGCAGCCCTT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTATTTCTGGAACAGATCTCGGATT		yes	yes
Pfat-4	WBGene00001396	*fat-4*	T13F2.1	GGGGACAACTTTGTATAGAAAAGTTGCCTTATAATAAAATAAAAAGTGTTTTCAAT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTGATATCACAGCGGT		yes	yes
Pacdh-2	WBGene00015894	*acdh-2*	C17C3.12	GGGGACAACTTTGTATAGAAAAGTTGCGGGTGCCTAATAATCAGAAA	GGGGACTGCTTTTTTGTACAAACTTGTCATTCTGAATCAGAACTGAGTAGTTA		yes	yes
Pfat-7	WBGene00001399	*fat-7*	F10D2.9	GGGGACAACTTTGTATAGAAAAGTTGCTTGATTGTAGATAAAAAAGTTTTG	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTACCGTTTGTTTTCTGAAAAGTATT		yes	yes
PM01B12.5	WBGene00019698	*M01B12.5*	M01B12.5	GGGGACAACTTTGTATAGAAAAGTTGCTTTTTTAAATAAAAACTCAGATAAAAT	GGGGACTGCTTTTTTGTACAAACTTGTCATGTTTGTTACCCACTGAA		yes	yes

Promoter name	*WB Gene ID*	*Public name*	*Sequence name*	[1] *GW-Forward Primer*	[2] *GW-Reverse Primer*	*Comments*	*Promoter cloned?*	*Y1H interactions after filtering?*
Pfat-3	WBGene00001395	*fat-3*	W08D2.4	GGGGACAACTTTGTATAGAAAAGTTGGATTACGATTAATTAATCAATTTATTTT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTGGTTGAGCGGCGGC		yes	yes
Pstr-47	WBGene00006112	*str-47*	F07C4.1	GGGGACAACTTTGTATAGAAAAGTTGGGAAAATCTTGATATTTTCGT	GGGGACTGCTTTTTTGTACAAACTTGTCATGGAGGATATGGAAGAA		yes	yes
PZK686.4	WBGene00022794	*ZK686.4*	ZK686.4	GGGGACAACTTTGTATAGAAAAGTTGGGAGAGTCGGAAGTTGATGT	GGGGACTGCTTTTTTGTACAAACTTGTCATCATGAGCCTCCGATT		yes	yes
PY57A10A.27	WBGene00013267	*Y57A10A.27*	Y57A10A.27	GGGGACAACTTTGTATAGAAAAGTTGGGGAATTTCGGCGGATTCT	GGGGACTGCTTTTTTGTACAAACTTGTCATAGCTGAAAAAATTGGATATTTATTCG		yes	yes
Phis-14	WBGene00001888	*his-14*	ZK131.8	GGGGACAACTTTGTATAGAAAAGTTGGGGCGGTTTGCTTAGTACGA	GGGGACTGCTTTTTTGTACAAACTTGTCATTGTTGACAATTGATGAAGACTC		yes	yes
Pgpd-3	WBGene00001685	*gpd-3*	K10B3.7	GGGGACAACTTTGTATAGAAAAGTTGGTTTTGATGTGCTTTTCATTTTGTACT	GGGGACTGCTTTTTTGTACAAACTTGTCATCTGTTTTGATTTCCTGGAAATGAAC		yes	yes
Pnhr-242	WBGene00022097	*nhr-242*	Y69A2AR.26	GGGGACAACTTTGTATAGAAAAGTTGTAACCATTCAAAACGGTGCAG	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTTTTGCAAGAGCTTCTCTCA		yes	yes
PF08F8.2	WBGene00017268	*F08F8.2*	F08F8.2	GGGGACAACTTTGTATAGAAAAGTTGTAATAATACTAATTCTTGTAATGTTC	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTGGTTTCCACCTGAAA		yes	yes
Pceh-60	WBGene00017690	*ceh-60*	F22A3.5	GGGGACAACTTTGTATAGAAAAGTTGTACCGATTTTTATATTTCTTATAAACT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTCAATAATCTTCGTGAATTT		yes	yes
PH32C10.3	WBGene00019257	*H32C10.3*	H32C10.3	GGGGACAACTTTGTATAGAAAAGTTGTCACATTTTGGATCAAAAAATAAATA	GGGGACTGCTTTTTTGTACAAACTTGTCATACATTGCATCAGTTTC		yes	yes
PT23F11.4	WBGene00011956	*T23F11.4*	T23F11.4	GGGGACAACTTTGTATAGAAAAGTTGTCCCTTGTTATTCCATATATAAA	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTAAGTCTGAAATAGAAAAATTTA		yes	yes
Pnhr-140	WBGene00016366	*nhr-140*	C33G8.9	GGGGACAACTTTGTATAGAAAAGTTGTGAAAGATGCAAACCTTTTCTCAGAGCTTCTA	GGGGACTGCTTTTTTGTACAAACTTGTCATCTGAAAATTGCAATTTTATTTTA		yes	yes
Pdop-3	WBGene00020506	*dop-3*	T14E8.3	GGGGACAACTTTGTATAGAAAAGTTGTGAGAGTCATTGTGATGGAAGG	GGGGACTGCTTTTTTGTACAAACTTGTCATTCTGATTTTAGACGACGGATT		yes	yes
Plpd-2	WBGene00003059	*lpd-2*	C48E7.3	GGGGACAACTTTGTATAGAAAAGTTGTGCAATGAAAACGTGTGGAAAGAA	GGGGACTGCTTTTTTGTACAAACTTGTCATTTTAGTTGAAGAAACTAGGCTCAAT		yes	yes

Promoter name	*WB Gene ID*	*Public name*	*Sequence name*	[1] *GW-Forward Primer*	[2] *GW-Reverse Primer*	*Comments*	*Promoter cloned?*	*Y1H interactions after filtering?*
Pptc-1	WBGene00004208	*ptc-1*	ZK675.1	GGGGACAACTTTGTATAGAAAAGTTGTTCAAAAAATAATTTAAATAGATAACTG	GGGGACTGCTTTTTTGTACAAACTTGTCATTGCTGCCGACTTGTCA		yes	yes
PC46E10.9	WBGene00016713	*C46E10.9*	C46E10 9	GGGGACAACTTTGTATAGAAAAGTTGTTGATAAGATTCTACTGCTTAAT	GGGGACTGCTTTTTTGTACAAACTTGTCATCGCATTCAATAACTCCCCT		yes	yes
Plin-1	WBGene00002990	*lin-1*	C37F5.1	GGGGACAACTTTGTATAGAAAAGTTGTTGTTAAAATGACAGTTTCAAATA	GGGGACTGCTTTTTTGTACAAACTTGTCATTGTAAACTGTAGAGTGTCGGC		yes	yes
Popt-2	WBGene00003877	*opt-2*	K04E7.2	GGGGACAACTTTGTATAGAAAAGTTGTTTCAAGGAATTCTTACCTG	GGGGACTGCTTTTTTGTACAAACTTGTCATAGTGGCGATACTGACGA		yes	yes
PC49C3.3	WBGene00008193	*C49C3.3*	C49C3.3	GGGGACAACTTTGTATAGAAAAGTTGTTTGGAGCATCAACTGAAATTTTTAG	GGGGACTGCTTTTTTGTACAAACTTGTCATAATTGACTGACCCTAGTTCC		yes	yes
Pdhs-25	WBGene00000988	*dhs-25*	F09E10.3	GGGGACAACTTTGTATAGAAAAGTTGTTTGTGAAAATGCCCAGCACTGTAT	GGGGACTGCTTTTTTGTACAAACTTGTCATAATTACCGACTGCTCGTTGCC		yes	yes
Papm-3	WBGene00000164	*apm-3*	F53H8.1	*promoterome*			yes	yes
Pfar-7	WBGene00001391	*far-7*	K01A2.2	*promoterome*			yes	yes
Pfat-2	WBGene00001394	*fat-2*	W02A2.1	*promoterome*			yes	yes
Phlh-15	WBGene00001959	*hlh-15*	C43H6.8	*promoterome*			yes	yes
Plbp-1	WBGene00002253	*lbp-1*	F40F4.3	*promoterome*			yes	yes
Pnhr-68	WBGene00003658	*nhr-68*	H12C20.3	*Deplancke et al.*			yes	yes
Pnhr-86	WBGene00003676	*nhr-86*	Y40B10A.8	GGGGACAACTTTGTATAGAAAAGTTGTAGTTTGGGATGGAAAACTAAATTGA	GGGGACTGCTTTTTTGTACAAACTTGTCATTGTTAGGCGGCGAACAAGG		yes	yes
Pnhr-137	WBGene00003727	*nhr-137*	C56E10.4	*promoterome*			yes	yes
Pref-2	WBGene00004335	*ref-2*	C47C12.3	*promoterome*			yes	yes
Psbp-1	WBGene00004735	*sbp-1*	Y47D3B.7	GGGGACAACTTTGTATAGAAAAGTTGCCAGGAGTTTTTGAAAAAATTCAAAATTCAAT	GGGGACTGCTTTTTTGTACAAACTTGTCATTCTGAAAAAAAAAAGTCAAATTTTGAG		yes	yes
PC36A4.9	WBGene00007969	*C36A4.9*	C36A4.9	*promoterome*			yes	yes
PF13D12.6	WBGene00008741	*F13D12.6*	F13D12.6	*promoterome*			yes	yes
Pacs-2	WBGene00009221	*acs-2*	F28F8.2	*promoterome*			yes	yes
PC15C7.5	WBGene00015791	*C15C7.5*	C15C7.5	*promoterome*			yes	yes
PC30F12.1	WBGene00016260	*C30F12.1*	C30F12.1	*promoterome*			yes	yes
Pnhr-178	WBGene00017510	*nhr-178*	F16B4.9	*promoterome*			yes	yes
Ptag-257	WBGene00018516	*tag-257*	F46G11 3	*promoterome*			yes	yes

Promoter name	*WB Gene ID*	*Public name*	*Sequence name*	[1] *GW-Forward Primer*	[2] *GW-Reverse Primer*	*Comments*	*Promoter cloned?*	*Y1H interactions after filtering?*
PF52C6.12	WBGene00018669	*F52C6.12*	F52C6.12	*promoterome*			yes	yes
Pcyp-35A5	WBGene00019473	*cyp-35A5*	K07C6.5	*promoterome*			yes	yes
Pegg-2	WBGene00019811	*egg-2*	R01H2.3	*promoterome*			yes	yes
Phlh-30	WBGene00020930	*hlh-30*	W02C12.3	GGGGACAACTTTGTATAGAAAAGTTGGTGTCTAAACTTTCTGATCGGGACCT	GGGGACTGCTTTTTTGTACAAACTTGTCATGAATGCTCTTATTCGCTG		yes	yes
Ppeb-1	WBGene00003968	*peb-1*	*T14F9.4*	*Deplancke et al.*			yes	yes
Pdaf-12b	WBGene00000908	*daf-12b*	F11A1.3b	*Vermeirssen et al.*			yes	yes
PK11D12.4	WBGene00019644	*cpt-4*	K11D12.4	GGGGACAACTTTGTATAGAAAAGTTGAATTAAAATAGTTGTTTTTTCGGAGGTG	GGGGACTGCTTTTTTGTACAAACTTGTCATTGTTTGTCTGAAAGAAAAACGTG	*promoter bait too self-active*	yes	N/A
PY40H7A.7	WBGene00012745	*Y40H7A.7*	Y40H7A.7	GGGGACAACTTTGTATAGAAAAGTTGTGTGCAGGTGGAGTACGGTAGA	GGGGACTGCTTTTTTGTACAAACTTGTCATTGCCACGCCGCTTTCGGCCTCCTT		yes	no
PC29F3.1	WBGene00001150	*ech-1*	C29F3.1	GGGGACAACTTTGTATAGAAAAGTTGTTGCGTGACACGTGGCATTT	GGGGACTGCTTTTTTGTACAAACTTGTCATCTGTAATAATGAGGTTTGAAGTAATATA	*cloning unsuccesful after multiple attempts*	no	N/A
PC34G6.4	WBGene00003996	*pgp-2*	C34G6.4	GGGGACAACTTTGTATAGAAAAGTTGCAAAAGAAGTAGAAACACCA	GGGGACTGCTTTTTTGTACAAACTTGTCATCAGTTAGATAATAATAGAAAATAAT	*cloning unsuccesful after multiple attempts*	no	N/A
PF08A8.2	WBGene00008565	*F08A8.2*	F08A8.2	GGGGACAACTTTGTATAGAAAAGTTGAAAGGAACTTTTCATAATCATCA	GGGGACTGCTTTTTTGTACAAACTTGTCATACTAAAACTGACAAAGTTTTAAAAATTTGA	*considered the promoter of F08A8.1 instead*	no	N/A
PF08A8.4	WBGene00008567	*F08A8.4*	P08A8.4			*considered the promoter of F08A8.1instead*	no	N/A
PF15H10.4	WBGene00008871	*tag-314*	F15H10.4	GGGGACAACTTTGTATAGAAAAGTTGAAAATTGATTTTTCTCTTTGTGG	GGGGACTGCTTTTTTGTACAAACTTGTCATTGTCGAAGAATGATACGAATCTG	*cloning unsuccesful after multiple attempts*	no	N/A
PK09H11.2	WBGene00006014	*srx-123*	K09H11.2			*pseudogene*	no	N/A
PR09B5.6	WBGene00019978	*hacd-1*	R09B5.6	GGGGACAACTTTGTATAGAAAAGTTGGTGACAAGTTGGTGACATGCTA	GGGGACTGCTTTTTTGTACAAACTTGTCATGTTTTGTGGGGGAAAATTTCTGA	*cloning unsuccesful after multiple attempts*	no	N/A
PY48G9A.10	WBGene00021703	*cpt-3*	Y48G9A.10	GGGGACAACTTTGTATAGAAAAGTTGTCCATAATGTTTATTTTTGTTTTTCTGTGCCT	GGGGACTGCTTTTTTGTACAAACTTGTCATTTCAACAATCTTAAACCGGTT	*promoter region too repetitive*	no	N/A

Promoter name	*WB Gene ID*	*Public name*	*Sequence name*	[1] *GW-Forward Primer*	[2] *GW-Reverse Primer*	*Comments*	*Promoter cloned?*	*Y1H interactions after filtering?*
[a] *attB4tail: GGGGACAACTTTGTATAGAAAAGTTG*								
[b] *attB1R tail: GGGGACTGCTTTTTTGTACAAACTTGTCAT*								

Table 2.2 List of preys used in Y1H matrix experiments

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
cfi-1	T23D8.8	WBGene00000476	ARID/BRIGHT	screen
F09G2.9	F09G2.9	WBGene00017317	AT Hook x3	other studies
hlh-15	C43H6.8	WBGene00001959	bHLH	Ashrafi et al
sbp-1	Y47D3B.7	WBGene00004735	bHLH	Ashrafi et al
hlh-30	W02C12.3	WBGene00020930	bHLH	Ashrafi et al/screen
ref-1	T01E8.2	WBGene00004334	bHLH - 2 domains	other studies
lpd-2	C48E7.3	WBGene00003059	bZIP	McKay et al
zip-2	K02F3.4	WBGene00019327	bZIP	screen
zip-4	Y44E3B.1	WBGene00021552	bZIP	other studies
ces-2	ZK909.4	WBGene00000469	bZIP	other studies
cey-2	F46F11.2	WBGene00000473	COLD BOX	screen
lin-28	F02E9.2	WBGene00003014	COLD BOX	other studies
T13C5.4	T13C5.4	WBGene00020485	HD	other studies
ceh-48	C17H12.9	WBGene00015934	HD - CUT	other studies
mab-5	C08C3.3	WBGene00003102	HD - HOX	screen
pal-1	C38D4.6	WBGene00003912	HD - HOX	screen
php-3	Y75B8A.1	WBGene00004024	HD - HOX	screen
lim-6	K03E6.1	WBGene00002988	HD - LIM	other studies

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
ceh-43	C28A5.4	WBGene00000463	HD - NK	screen
tab-1	F31E8.3	WBGene00006380	HD - NK	screen
ceh-23	ZK652.5	WBGene00000446	HD - NK	screen
mls-2	C39E6.4	WBGene00003377	HD - NK	other studies
ceh-6	K02B12.1	WBGene00000431	HD - POU	screen
ceh-10	W03A3.1	WBGene00000435	HD - PRD	screen
unc-30	B0564.10	WBGene00006766	HD - PRD	other studies
C09G12.1	C09G12.1	WBGene00015651	HD - PRD	other studies
dsc-1	C18B12.3	WBGene00001096	HD - PRD	other studies
ceh-36	C37E2.4	WBGene00000457	HD - PRD	other studies
ceh-37	C37E2.5	WBGene00000458	HD - PRD	other studies
ceh-17	D1007.1	WBGene00000440	HD - PRD	other studies
unc-42	F58E6.10	WBGene00006778	HD - PRD	other studies
alr-1	R08B4.2	WBGene00044330	HD - PRD	other studies
unc-4	UNC-4	WBGene00006744	HD - PRD	other studies
ttx-1	Y113G7A.6	WBGene00006652	HD - PRD	other studies
ceh-45	ZK993.1	WBGene00022837	HD - PRD	other studies
ceh-8	ZK265.4	WBGene00000433	HD - PRD - 2 domains	screen
Y53C12C.1	Y53C12C.1	WBGene00013147	HD - PRD, Paired Domain - CPAX	other studies

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
pax-3	F27E5.2	WBGene00003939	HD - PRD, Paired Domain - FULL	other studies
vab-3	F14F3.1	WBGene00006870	HD - PRD, Paired Domain - FULL	screen
ceh-26	K12H4.1	WBGene00000448	HD - PROX	screen
ceh-20	F31E3.1	WBGene00000443	HD - TALE	screen
ceh-60	F22A3.5	WBGene00017690	HD -TALE	Ashrafi et al
daf-3	F25E2.5	WBGene00000899	MH1	screen
sma-3	R13F6.9	WBGene00004857	MH1	screen
B0261.1	B0261.1	WBGene00015091	MYB	screen
B0238.11	B0238.11	WBGene00015075	Novel	screen
C32D5.1	C32D5.1	WBGene00016310	Novel	screen
prx-5	C34C6.6	WBGene00004194	Novel	screen
C44F1.1	C44F1.1	WBGene00008091	Novel	screen
F08G12.3	F08G12.3	WBGene00008587	Novel	screen
H02I12.5	H02I12.5	WBGene00010353	Novel	screen
lin-54	JC8.6	WBGene00003037	Novel	screen
T20F10.2	T20F10.2	WBGene00011864	Novel	screen
Y17G7B.20	Y17G7B.20	WBGene00012471	Novel	screen
Y17G9B.9	Y17G9B.9	WBGene00021206	Novel	screen
Y38C9A.1	Y38C9A.1	WBGene00021411	Novel	screen

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
Y53G8AM.8	Y53G8AM.8	WBGene00021808	Novel	screen
Y62E10A.14	Y62E10A.14	WBGene00013380	Novel	screen
ZC204.12	ZC204.12	WBGene00022562	Novel	screen
C34C12.4	C34C12.4	WBGene00007923	Novel	other studies
C39F7.2	C39F7.2	WBGene00016539	Novel	other studies
ftt-2	F52D10.3	WBGene00001502	Novel	other studies
sdz-24	K07E8.3	WBGene00019495	Novel	other studies
T06A10.4	T06A10.4	WBGene00020287	Novel	other studies
Y52E8A.2	Y52E8A.2	WBGene00021795	Novel	other studies
cec-1	ZK1236.2	WBGene00000414	Novel	other studies
pax-2	K06B9.5	WBGene00003938	Paired Domain - FULL	other studies
C09G9.7	C09G9.7	WBGene00007496	Paired Domain - NPAX	Ashrafi et al
F26H9.2	F26H9.2	WBGene00009174	RPEL - 2 domains	other studies
tbx-8	T07C4.2	WBGene00006545	T-box	screen
egl-44	F28B12.2	WBGene00001208	TEA/ATTS	screen
Y65B4BR.5	Y65B4BR.5	WBGene00022042	UNKNOWN	screen
mig-5	T05C12.6	WBGene00003241	WH	screen
T28D9.9	T28D9.9	WBGene00020898	WH	screen
lin-1	C37F5.1	WBGene00002990	WH - ETS	Ashrafi et al

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
fkh-2	T14G12.4	WBGene00001434	WH - Fork Head	screen
Y47D3B.9	Y47D3B.9	WBGene00012943	ZF - BED	screen
F55B11.4	F55B11.4	WBGene00010086	ZF - C2H2 - 1 finger	Ashrafi et al
T09F3.1	T09F3.1	WBGene00011661	ZF - C2H2 - 1 finger	Ashrafi et al
T23F11.4	T23F11.4	WBGene00011956	ZF - C2H2 - 1 finger	Ashrafi et al
ZK686.4	ZK686.4	WBGene00022794	ZF - C2H2 - 1 finger	Ashrafi et al
ztf-3	C53D5.4	WBGene00016905	ZF - C2H2 - 1 finger	screen
lir-1	F18A1.3	WBGene00003044	ZF - C2H2 - 1 finger	screen
rabs-5	Y42H9AR.3	WBGene00021538	ZF - C2H2 - 1 finger	other studies
ztf-6	W06H12.1	WBGene00012317	ZF - C2H2 - 2 fingers	screen
ref-2	C47C12.3	WBGene00004335	ZF - C2H2 - 3 fingers	Ashrafi et al
ztf-2	F13G3.1	WBGene00008762	ZF - C2H2 - 3 fingers	direct transformation
die-1	C18D1.1	WBGene00000995	ZF - C2H2 - 3 fingers	screen
ztf-8	ZC395.8	WBGene00022598	ZF - C2H2 - 3 fingers	screen
sptf-3	Y40B1A.4	WBGene00012735	ZF - C2H2 - 3 fingers	other studies
C46E10.9	C46E10.9	WBGene00016713	ZF - C2H2 - 4 fingers	Ashrafi et al
ces-1	F43G9.11	WBGene00000468	ZF - C2H2 - 4 fingers	screen
ztf-1	F54F2.5	WBGene00018833	ZF - C2H2 - 4 fingers	screen
F10B5.3	F10B5.3	WBGene00008640	ZF - C2H2 - 4 fingers	other studies

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
syd-9	ZK867.1	WBGene00044068	ZF - C2H2 - 4 fingers	other studies
lsy-2	F49H12.1	WBGene00003087	ZF - C2H2 - 5 fingers	screen
pag-3	F45B8.4	WBGene00003909	ZF - C2H2 - 5 fingers	other studies
ztf-11	F52F12.6	WBGene00009939	ZF - C2HC 2 fingers	other studies
F38B7.1	F38B7.1	WBGene00009532	ZF - CCCH - 2 domains	screen
pie-1	Y49E10.14	WBGene00004027	ZF - CCCH - 2 domains	screen
Y57G11C.25	Y57G11C.25	WBGene00013319	ZF - CCCH - 2 domains	screen
mex-5	W02A2.7	WBGene00003230	ZF - CCCH - 2 domains	other studies
H32C10.3	H32C10.3	WBGene00019257	ZF - DHHC	Ashrafi et al
K08B12.2	K08B12.2	WBGene00019521	ZF - DM	other studies
flh-2	C26E6.2	WBGene00016138	ZF - FLYWCH	screen
flh-1	Y11D7A.12	WBGene00012435	ZF - FLYWCH	screen
elt-7	C18G1.2	WBGene00015981	ZF - GATA	screen
elt-2	C33D3.1	WBGene00001250	ZF - GATA	screen
elt-6	F52C12.5	WBGene00001253	ZF - GATA	screen
egl-18	F55A8.1	WBGene00001186	ZF - GATA	screen
elt-3	K02B9.4	WBGene00001251	ZF - GATA	screen
elt-4	C39B10.6	WBGene00001252	ZF - GATA	other studies
end-1	F58E10.2	WBGene00001310	ZF - GATA	other studies

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
end-3	F58E10.5	WBGene00001311	ZF - GATA	other studies
med-2	K04C2.6	WBGene00003181	ZF - GATA	other studies
elt-1	W09C2.1	WBGene00001249	ZF - GATA - 2 domains	screen
lin-40	T27C4.4	WBGene00003025	ZF - GATA, MYB	other studies
Y74C9A.4	Y74C9A.4	WBGene00022278	ZF - GATA, MYB (2x)	other studies
nhr-140	C33G8.9	WBGene00016366	ZF - NHR	Ashrafi et al
nhr-137	C56E10.4	WBGene00003727	ZF - NHR	Ashrafi et al
daf-12b	F11A1.3b	WBGene00000908	ZF - NHR	Ashrafi et al
nhr-178	F16B4.9	WBGene00017510	ZF - NHR	Ashrafi et al
nhr-8	F33D4.1	WBGene00003607	ZF - NHR	Ashrafi et al
nhr-68	H12C20.3	WBGene00003658	ZF - NHR	Ashrafi et al
nhr-88	K08A2.5	WBGene00003678	ZF - NHR	Ashrafi et al
nhr-49	K10C3.6	WBGene00003639	ZF - NHR	Ashrafi et al
nhr-12	R04B5.4	WBGene00003611	ZF - NHR	direct transformation
nhr-273	T12C9.1	WBGene00020460	ZF - NHR	direct transformation
nhr-114	Y45G5AM.1	WBGene00003704	ZF - NHR	direct transformation
nhr-112	Y70C5C.6	WBGene00003702	ZF - NHR	direct transformation
nhr-10	B0280.8	WBGene00003609	ZF - NHR	screen
nhr-23	C01H6.5	WBGene00003622	ZF - NHR	screen

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
nhr-67	C08F8.8	WBGene00003657	ZF - NHR	screen
nhr-28	C11G6.4	WBGene00003624	ZF - NHR	screen
nhr-136	C13C4.3	WBGene00003726	ZF - NHR	screen
nhr-43	C29E6.5	WBGene00003633	ZF - NHR	screen
nhr-2	C32F10.6	WBGene00003601	ZF - NHR	screen
nhr-6	C48D5.1	WBGene00003605	ZF - NHR	screen
nhr-179	F16B4.11	WBGene00017512	ZF - NHR	screen
nhr-45	F16H11.5	WBGene00003635	ZF - NHR	screen
sex-1	F44A6.2	WBGene00004786	ZF - NHR	screen
nhr-111	F44G3.9	WBGene00003701	ZF - NHR	screen
nhr-34	F58G6.5	WBGene00003627	ZF - NHR	screen
nhr-207	R07B7.14	WBGene00011098	ZF - NHR	screen
nhr-102	T06C12.6	WBGene00003692	ZF - NHR	screen
nhr-66	T09A12.4	WBGene00003656	ZF - NHR	screen
odr-7	T18D3.2	WBGene00003854	ZF - NHR	screen
nhr-79	T26H2.9	WBGene00003669	ZF - NHR	screen
nhr-91	Y15E3A.1	WBGene00003681	ZF - NHR	screen
nhr-65	Y17D7A.3	WBGene00003655	ZF - NHR	screen
nhr-86	Y40B10A.8	WBGene00003676	ZF - NHR	screen

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
nhr-237	Y46H3D.6	WBGene00021610	ZF - NHR	screen
nhr-70	Y51A2D.17	WBGene00003660	ZF - NHR	screen
nhr-13	Y5H2B.2	WBGene00003612	ZF - NHR	screen
nhr-246	ZK1037.4	WBGene00014192	ZF - NHR	screen
nhr-35	C07A12.3	WBGene00003628	ZF - NHR	other studies
nhr-46	C45E5.6	WBGene00003636	ZF - NHR	other studies
nhr-141	F25E5.6	WBGene00017787	ZF - NHR	other studies
nhr-4	F32B6.1	WBGene00003603	ZF - NHR	other studies
nhr-142	F44E7.8	WBGene00018430	ZF - NHR	other studies
fax-1	F56E3.4	WBGene00001400	ZF - NHR	other studies
nhr-3	H01A20.1	WBGene00003602	ZF - NHR	other studies
nhr-22	K06A1.4	WBGene00003621	ZF - NHR	other studies
nhr-98	M02H5.6	WBGene00003688	ZF - NHR	other studies
nhr-123	M02H5.7	WBGene00003713	ZF - NHR	other studies
nhr-1	R09G11.2	WBGene00003600	ZF - NHR	other studies
nhr-14	T01B10.4	WBGene00003613	ZF - NHR	other studies
nhr-84	T06C12.7	WBGene00003674	ZF - NHR	other studies
nhr-109	T12C9.5	WBGene00003699	ZF - NHR	other studies
nhr-69	T23H4.2	WBGene00003659	ZF - NHR	other studies

Public/CGC name	*Sequence name*	*Wormbase ID*	*DNA binding domain (wTF2.1)*	*source*
nhr-61	W01D2.2	WBGene00003651	ZF - NHR	other studies
nhr-41	Y104H12A.1	WBGene00022423	ZF - NHR	other studies
nhr-234	Y38E10A.18	WBGene00012596	ZF - NHR	other studies
peb-1	T14F9.4	WBGene00003968	ZF - FLYWCH	Ashrafi et al

Table 2.3 High confidence promoter-TF interactions (normalized score ≥ 5.0)

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pacs-11	CEH-10	1	1	1	1	1	1	0	0	0	5.0
Pacs-2	MIG-5	1	1	1	0	1	1	0	N/A	N/A	5.0
PC06E7.3	PHP-3	1	1	1	1	1	1	0	0	0	5.0
PC49C3.3	Y57G11C.25	1	1	1	1	1	1	0	0	0	5.0
Pceh-60	NHR-6	1	1	1	1	1	1	0	0	0	5.0
PF52C6.12	C32D5.1	0	1	0	1	0	0	0	1	N/A	5.0
PF52C6.12	Y17G7B.20	0	1	0	1	0	0	0	1	N/A	5.0
PF55B11.4	CFI-1	1	1	1	1	1	1	0	0	0	5.0
Pfar-7	NHR-6	1	1	1	1	1	1	0	0	0	5.0
Pfat-2	EGL-44	1	1	1	0	0	1	0	1	0	5.0
Phlh-30	C32D5.1	0	1	0	1	0	0	0	1	N/A	5.0
Plin-1	DIE-1	0	1	1	1	1	1	0	0	N/A	5.0
Plin-1	H02I12.5	0	1	0	0	1	1	0	1	N/A	5.0
Plin-1	PAL-1	0	1	1	1	1	1	0	0	N/A	5.0
Pnhr-178	TBX-33	1	1	1	1	0	0	0	N/A	N/A	5.0
Pnhr-49	H02I12.5	0	1	0	0	1	1	0	1	N/A	5.0
Popt-2	F26H9.2	1	1	1	0	1	1	0	N/A	N/A	5.0
Popt-2	T20F10.2	1	1	0	0	1	1	0	1	0	5.0
Pptc-1	C32D5.1	0	1	0	1	0	0	0	1	N/A	5.0
Pptc-1	LIN-54	0	1	0	1	0	0	0	1	N/A	5.0
Psbp-1	C32D5.1	0	1	0	1	0	0	0	1	N/A	5.0
Psbp-1	DIE-1	0	1	1	1	1	1	0	0	N/A	5.0
Psbp-1	LIN-54	0	1	0	1	0	0	0	1	N/A	5.0
Psbp-1	PHP-3	0	1	1	1	1	1	0	0	N/A	5.0
PT23F11.4	C32D5.1	0	1	0	1	0	0	0	1	N/A	5.0
PT23F11.4	LIN-54	0	1	0	1	0	0	0	1	N/A	5.0

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pacs-2	TAB-1	1	1	1	0	1	1	0	1	0	5.7
Papm-3	NHR-23	1	1	1	1	1	1	1	0	0	5.7
PC30F12.1	ZTF-8	1	1	1	0	1	1	0	0	1	5.7
PC36A4.9	TAB-1	1	1	1	0	1	1	0	1	0	5.7
PC36A4.9	ZTF-8	1	1	1	0	1	1	0	0	1	5.7
Pceh-60	EGL-44	1	1	1	0	1	1	0	1	0	5.7
Pcyp-35A5	TAB-1	1	1	1	0	1	1	0	1	0	5.7
Pelo-2	ZTF-1	1	1	1	1	0	0	0	0	1	5.7
Pfar-7	TAB-1	1	1	1	0	1	1	0	1	0	5.7
Pfat-7	TAB-1	1	1	1	0	1	1	0	1	0	5.7
Phis-14	TAB-1	1	1	1	0	1	1	0	1	0	5.7
Phlh-15	H02I12.5	1	1	0	0	1	1	1	1	0	5.7
Phlh-15	UNC-42	1	1	1	1	0	0	0	1	0	5.7
Phlh-30	FLH-2	0	1	1	0	0	0	0	1	1	5.7
Plpd-2	ZC204.12	1	1	0	1	0	1	0	0	1	5.7
Pmdt-15	NHR-111	1	1	1	0	1	1	0	1	0	5.7
Pmdt-15	TAB-1	1	1	1	0	1	1	0	1	0	5.7
Pmdt-15	ZTF-8	1	1	1	0	1	1	0	0	1	5.7
Pnhr-25	T20F10.2	1	1	0	0	0	0	0	1	1	5.7
Pnhr-8	EGL-44	1	1	1	0	1	1	0	1	0	5.7
Pnhr-8	TAB-1	1	1	1	0	1	1	0	1	0	5.7
Popt-2	EGL-44	1	1	1	0	1	1	0	1	0	5.7
Popt-2	TAB-1	1	1	1	0	1	1	0	1	0	5.7
Psrj-34	EGL-44	1	1	1	0	1	1	0	1	0	5.7
Psrj-34	TAB-1	1	1	1	0	1	1	0	1	0	5.7
Pstr-47	ZTF-8	1	1	1	0	1	1	0	0	1	5.7
PT09F3.1	TAB-1	1	1	1	0	1	1	0	1	0	5.7

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Ptag-257	PAL-1	1	1	1	1	1	1	1	0	0	5.7
PZK686.4	TAB-1	1	1	1	0	1	1	0	1	0	5.7
PC46E10.9	CEH-48	0	1	1	1	0	0	0	1	N/A	5.8
PC46E10.9	LSY-2	0	1	1	1	0	0	0	1	N/A	5.8
PC46E10.9	ODR-7	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	CEH-48	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	DAF-3	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	DIE-1	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	END-3	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	K08B12.2	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	LIR-1	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	NHR-1	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	NHR-43	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	PAG-3	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	PHP-3	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	PIE-1	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	T09F3.1	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	TBX-8	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	VAB-3	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	HLH-30	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	FLH-1	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	Y47D3B.9	0	1	1	1	0	0	0	1	N/A	5.8
PF52C6.12	ZTF-1	0	1	1	1	0	0	0	1	N/A	5.8
Pfat-3	LSY-2	0	1	1	1	0	0	0	1	N/A	5.8
Pgei-7	CEH-10	0	1	1	1	0	0	0	1	N/A	5.8
Pgei-7	NHR-2	0	1	1	1	0	0	0	1	N/A	5.8
Pgei-7	ODR-7	0	1	1	1	0	0	0	1	N/A	5.8

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pgei-7	ZTF-1	0	1	1	1	0	0	0	1	N/A	5.8
Phlh-30	DAF-3	0	1	1	1	0	0	0	1	N/A	5.8
Phlh-30	ODR-7	0	1	1	1	0	0	0	1	N/A	5.8
Phlh-30	TAB-1	0	1	1	0	1	1	0	1	N/A	5.8
Plin-1	NHR-2	0	1	1	1	0	0	0	1	N/A	5.8
Plin-1	NHR-91	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	CEH-10	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	CES-1	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	DSC-1	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	NHR-273	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	NHR-41	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	NHR-86	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	NHR-91	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	PAL-1	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	TBX-8	0	1	1	1	0	0	0	1	N/A	5.8
Pnhr-49	ZTF-1	0	1	1	1	0	0	0	1	N/A	5.8
Pptc-1	NHR-41	0	1	1	1	0	0	0	1	N/A	5.8
Pptc-1	ODR-7	0	1	1	1	0	0	0	1	N/A	5.8
Pptc-1	TBX-8	0	1	1	1	0	0	0	1	N/A	5.8
Pptc-1	Y47D3B.9	0	1	1	1	0	0	0	1	N/A	5.8
Pptc-1	Y53C12C.1	0	1	1	1	0	0	0	1	N/A	5.8
Psbp-1	CEH-10	0	1	1	1	0	0	0	1	N/A	5.8
Psbp-1	ELT-6	0	1	1	1	0	0	0	1	N/A	5.8
Psbp-1	NHR-102	0	1	1	1	0	0	0	1	N/A	5.8
Psbp-1	NHR-111	0	1	1	0	1	1	0	1	N/A	5.8
Psbp-1	NHR-67	0	1	1	1	0	0	0	1	N/A	5.8
Psbp-1	NHR-91	0	1	1	1	0	0	0	1	N/A	5.8

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Psbp-1	ODR-7	0	1	1	1	0	0	0	1	N/A	5.8
PT23F11.4	CEH-20	0	1	1	1	0	0	0	1	N/A	5.8
PT23F11.4	DAF-3	0	1	1	1	0	0	0	1	N/A	5.8
PT23F11.4	DSC-1	0	1	1	1	0	0	0	1	N/A	5.8
PT23F11.4	LIN-1	0	1	1	1	0	0	0	1	N/A	5.8
PT23F11.4	NHR-234	0	1	1	1	0	0	0	1	N/A	5.8
PT23F11.4	ODR-7	0	1	1	1	0	0	0	1	N/A	5.8
PT23F11.4	UNC-42	0	1	1	1	0	0	0	1	N/A	5.8
PT23F11.4	ZTF-6	0	1	1	1	0	0	0	1	N/A	5.8
PZK593.3	CES-2	0	1	1	1	0	0	0	1	N/A	5.8
PZK593.3	NHR-41	0	1	1	1	0	0	0	1	N/A	5.8
Ppeb-1	NHR-1	0	1	1	1	0	0	0	1	N/A	5.8
Ppeb-1	NHR-2	0	1	1	1	0	0	0	1	N/A	5.8
Ppeb-1	NHR-41	0	1	1	1	0	0	0	1	N/A	5.8
Ppeb-1	ELT-1	0	1	1	1	0	0	0	1	N/A	5.8
Ppeb-1	ELT-2	0	1	1	1	0	0	0	1	N/A	5.8
Ppeb-1	ELT-3	0	1	1	1	0	0	0	1	N/A	5.8
Plpd-2	Y37D8A.11	1	1	0	1	1	1	0	N/A	N/A	6.0
Pacs-11	EGL-44	1	1	1	0	0	0	0	1	1	6.4
PC15C7.5	TAB-1	1	1	1	0	0	0	0	1	1	6.4
PC36A4.9	ZTF-1	1	1	1	1	0	1	0	0	1	6.4
PC49C3.3	TAB-1	1	1	1	0	0	0	0	1	1	6.4
Pceh-60	FLH-2	1	1	1	0	0	0	0	1	1	6.4
Pech-6	ZTF-8	1	1	1	0	0	0	0	1	1	6.4
Pegg-2	NHR-111	1	1	1	0	0	0	0	1	1	6.4
Pegg-2	ZC204.12	1	1	0	1	1	1	0	0	1	6.4
Pelo-2	DIE-1	1	1	1	1	0	1	0	0	1	6.4

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
PF55B11.4	LIN-54	1	1	0	1	1	1	0	0	1	6.4
PF59F5.2	NHR-111	1	1	1	0	0	0	0	1	1	6.4
Pfat-2	MIG-5	1	1	1	0	0	0	0	1	1	6.4
Pfat-2	ZTF-8	1	1	1	0	0	0	0	1	1	6.4
Pfat-4	EGL-44	1	1	1	0	0	0	0	1	1	6.4
Pgpd-3	NHR-111	1	1	1	0	0	0	0	1	1	6.4
Plbp-8	TAB-1	1	1	1	0	0	0	0	1	1	6.4
PM01B12.5	EGL-44	1	1	1	0	0	0	0	1	1	6.4
PM01B12.5	ZC204.12	1	1	0	1	1	1	0	1	0	6.4
Pnhr-140	LIN-54	1	1	0	1	1	1	0	1	0	6.4
Pnhr-178	NHR-111	1	1	1	0	0	0	0	1	1	6.4
Pnhr-25	F26H9.2	1	1	1	0	0	0	0	1	1	6.4
Pnhr-68	NHR-111	1	1	1	0	0	0	0	1	1	6.4
Pnhr-8	NHR-111	1	1	1	0	0	0	0	1	1	6.4
Pnhr-86	MAB-5	1	1	1	1	0	1	0	1	0	6.4
Popt-2	MIG-5	1	1	1	0	0	0	0	1	1	6.4
Pptc-1	ZTF-1	0	1	1	1	0	1	0	1	N/A	6.7
PF59F5.2	ZAG-1	1	1	1	1	1	1	0	N/A	N/A	7.0
Pacdh-2	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pacs-11	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pacs-11	LIN-54	1	1	0	1	0	0	0	1	1	7.1
Pacs-11	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pacs-11	ZC204.12	1	1	0	1	1	1	1	0	1	7.1
Papm-3	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Papm-3	LIN-54	1	1	0	1	0	0	0	1	1	7.1
Papm-3	ZC204.12	1	1	0	1	0	0	0	1	1	7.1
PC06E7.3	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
PC09G9.7	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
PC15C7.5	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
PC30F12.1	EGL-44	1	1	1	0	0	1	0	1	1	7.1
PC36A4.9	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
PC36A4.9	PHP-3	1	1	1	1	1	1	0	1	0	7.1
PC49C3.3	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
PC49C3.3	DIE-1	1	1	1	1	1	1	0	0	1	7.1
PC49C3.3	LIN-54	1	1	0	1	0	0	0	1	1	7.1
PC49C3.3	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pceh-60	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pcyp-35A3	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pcyp-35A3	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pcyp-35A3	ZC204.12	1	1	0	1	1	1	1	0	1	7.1
Pdaf-12b	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pech-6	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pech-6	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pegg-2	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pegg-2	PIE-1	1	1	1	1	1	1	0	1	0	7.1
Pegg-2	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pelo-2	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pepn-1	F08G12.3	1	1	0	1	0	0	0	1	1	7.1
Pepn-1	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
PF08F8.2	PIE-1	1	1	1	1	1	1	0	1	0	7.1
PF13D12.6	LIN-54	1	1	0	1	0	0	0	1	1	7.1
PF52C6.12	CEH-8	0	1	1	1	0	0	0	1	1	7.1
PF52C6.12	ELT-1	0	1	1	1	0	0	0	1	1	7.1
PF52C6.12	ELT-2	0	1	1	1	0	0	0	1	1	7.1

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
PF52C6.12	NHR-111	0	1	1	0	1	1	0	1	1	7.1
PF52C6.12	ODR-7	0	1	1	1	0	0	0	1	1	7.1
PF55B11.4	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
PF55B11.4	DIE-1	1	1	1	1	1	1	0	0	1	7.1
PF55B11.4	ELT-3	1	1	1	1	1	1	0	1	0	7.1
PF55B11.4	MAB-5	1	1	1	1	1	1	0	1	0	7.1
PF55B11.4	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
PF59F5.2	F08G12.3	1	1	0	1	0	0	0	1	1	7.1
Pfar-7	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pfat-2	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pfat-2	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pfat-2	ZC204.12	1	1	0	1	0	0	0	1	1	7.1
Pfat-4	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pfat-7	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pgpd-3	F08G12.3	1	1	0	1	0	0	0	1	1	7.1
Phis-14	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Phlh-15	CEH-8	1	1	1	1	1	1	0	1	0	7.1
Phlh-15	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Phlh-30	CEH-48	0	1	1	1	0	0	0	1	1	7.1
Plbp-1	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Plbp-1	PAX-2	1	1	1	1	1	1	0	1	0	7.1
Plbp-8	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Plpd-2	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
PM01B12.5	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
PM01B12.5	ZTF-6	1	1	1	1	1	1	0	1	0	7.1
Pmdt-15	F08G12.3	1	1	0	1	0	0	0	1	1	7.1
Pmdt-15	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pnhr-137	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pnhr-178	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pnhr-178	PRX-5	1	1	0	0	1	1	0	1	1	7.1
Pnhr-178	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pnhr-25	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pnhr-68	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pnhr-8	ODR-7	1	1	1	1	1	1	0	1	0	7.1
Pnhr-8	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pnhr-86	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
Pnhr-86	Y17G7B.20	1	1	0	1	0	0	0	1	1	7.1
Pnhr-86	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pnhr-86	ZC204.12	1	1	0	1	1	1	1	1	0	7.1
Pnhr-88	ZTF-1	1	1	1	1	1	1	0	1	0	7.1
Pref-2	TBX-8	1	1	1	1	1	1	0	0	1	7.1
Pref-2	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pstr-47	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
PT09F3.1	LIN-54	1	1	0	1	0	0	0	1	1	7.1
PT09F3.1	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
PY57A10A.27	C32D5.1	1	1	0	1	0	0	0	1	1	7.1
PY57A10A.27	Y17G7B.20	1	1	0	1	1	1	1	1	0	7.1
PY57A10A.27	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
PZK686.4	Y38C9A.1	1	1	0	1	0	0	0	1	1	7.1
Pnhr-49	PHP-3	0	1	1	1	1	1	0	1	N/A	7.5
PT23F11.4	DIE-1	0	1	1	1	1	1	0	1	N/A	7.5
PZK593.3	SEX-1	0	1	1	1	1	1	0	1	N/A	7.5
Pacdh-1	CEH-8	1	1	1	1	0	0	0	1	1	7.9
Pacdh-1	NHR-10	1	1	1	1	0	0	0	1	1	7.9

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pacdh-1	UNC-42	1	1	1	1	0	0	0	1	1	7.9
Pacdh-2	ELT-2	1	1	1	1	0	0	0	1	1	7.9
Pacdh-2	NHR-1	1	1	1	1	0	0	0	1	1	7.9
Pacs-11	CEH-20	1	1	1	1	0	0	0	1	1	7.9
Pacs-11	CEH-6	1	1	1	1	0	0	0	1	1	7.9
Pacs-11	DIE-1	1	1	1	1	0	0	0	1	1	7.9
Pacs-11	UNC-42	1	1	1	1	0	0	0	1	1	7.9
Pacs-11	Y47D3B.9	1	1	1	1	0	0	0	1	1	7.9
Pacs-11	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pacs-11	ZTF-2	1	1	1	1	0	0	0	1	1	7.9
Papm-3	CEH-10	1	1	1	1	0	0	0	1	1	7.9
Papm-3	CEH-20	1	1	1	1	0	0	0	1	1	7.9
Papm-3	CEH-8	1	1	1	1	0	0	0	1	1	7.9
Papm-3	DIE-1	1	1	1	1	0	0	0	1	1	7.9
Papm-3	EKL-2	1	1	1	1	0	0	0	1	1	7.9
Papm-3	NHR-1	1	1	1	1	0	0	0	1	1	7.9
Papm-3	NHR-2	1	1	1	1	0	0	0	1	1	7.9
Papm-3	SEX-1	1	1	1	1	0	0	0	1	1	7.9
Papm-3	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
PC06E7.3	MAB-5	1	1	1	1	1	1	1	1	0	7.9
PC06E7.3	NHR-6	1	1	1	1	0	0	0	1	1	7.9
PC15C7.5	CEH-10	1	1	1	1	0	0	0	1	1	7.9
PC15C7.5	CEH-17	1	1	1	1	0	0	0	1	1	7.9
PC15C7.5	CEH-8	1	1	1	1	0	0	0	1	1	7.9
PC15C7.5	DIE-1	1	1	1	1	0	0	0	1	1	7.9
PC15C7.5	DSC-1	1	1	1	1	0	0	0	1	1	7.9
PC15C7.5	NHR-273	1	1	1	1	0	0	0	1	1	7.9

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
PC15C7.5	UNC-42	1	1	1	1	0	0	0	1	1	7.9
PC15C7.5	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
PC36A4.9	MLS-2	1	1	1	1	0	0	0	1	1	7.9
PC36A4.9	NHR-1	1	1	1	1	0	0	0	1	1	7.9
PC36A4.9	NHR-69	1	1	1	1	0	0	0	1	1	7.9
PC36A4.9	NHR-86	1	1	1	1	0	0	0	1	1	7.9
PC36A4.9	PAL-1	1	1	1	1	1	1	1	1	0	7.9
PC36A4.9	UNC-42	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	ALR-1	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	CEH-23	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	CEH-43	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	CEH-6	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	CEH-8	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	DSC-1	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	MAB-5	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	NHR-6	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	T13C5.4	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	TBX-8	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	UNC-42	1	1	1	1	0	0	0	1	1	7.9
PC49C3.3	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pceh-60	CEH-8	1	1	1	1	0	0	0	1	1	7.9
Pceh-60	NHR-86	1	1	1	1	0	0	0	1	1	7.9
Pceh-60	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pcyp-35A3	ODR-7	1	1	1	1	0	0	0	1	1	7.9
Pcyp-35A3	Y47D3B.9	1	1	1	1	0	0	0	1	1	7.9
Pcyp-35A3	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pcyp-35A5	MLS-2	1	1	1	1	0	0	0	1	1	7.9

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pdaf-12b	DIE-1	1	1	1	1	1	1	1	1	0	7.9
Pdhs-25	NHR-1	1	1	1	1	0	0	0	1	1	7.9
Pdhs-25	NHR-10	1	1	1	1	0	0	0	1	1	7.9
Pdhs-25	NHR-111	1	1	1	0	1	1	0	1	1	7.9
Pdhs-25	NHR-2	1	1	1	1	0	0	0	1	1	7.9
Pdhs-25	NHR-41	1	1	1	1	0	0	0	1	1	7.9
Pdhs-25	SEX-1	1	1	1	1	0	0	0	1	1	7.9
Pdhs-25	TBX-8	1	1	1	1	0	0	0	1	1	7.9
Pdop-3	TBX-8	1	1	1	1	0	0	0	1	1	7.9
Pech-6	CEH-10	1	1	1	1	0	0	0	1	1	7.9
Pech-6	CEH-17	1	1	1	1	0	0	0	1	1	7.9
Pech-6	CEH-6	1	1	1	1	0	0	0	1	1	7.9
Pech-6	CEH-8	1	1	1	1	0	0	0	1	1	7.9
Pech-6	LIN-54	1	1	0	1	0	1	0	1	1	7.9
Pech-6	NHR-43	1	1	1	1	0	0	0	1	1	7.9
Pech-6	ODR-7	1	1	1	1	0	0	0	1	1	7.9
Pech-6	T13C5.4	1	1	1	1	0	0	0	1	1	7.9
Pech-6	UNC-42	1	1	1	1	0	0	0	1	1	7.9
Pech-6	VAB-3	1	1	1	1	0	0	0	1	1	7.9
Pegg-2	CEH-8	1	1	1	1	0	0	0	1	1	7.9
Pegg-2	NHR-86	1	1	1	1	0	0	0	1	1	7.9
Pegg-2	TBX-8	1	1	1	1	0	0	0	1	1	7.9
Pegg-2	UNC-42	1	1	1	1	0	0	0	1	1	7.9
Pegg-2	ZTF-2	1	1	1	1	0	0	0	1	1	7.9
Pelo-2	CEH-10	1	1	1	1	0	0	0	1	1	7.9
Pelo-2	TBX-8	1	1	1	1	0	0	0	1	1	7.9
Pepn-1	CEH-8	1	1	1	1	0	0	0	1	1	7.9

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
PF13D12.6	CEH-10	1	1	1	1	0	0	0	1	1	7.9
PF13D12.6	CEH-17	1	1	1	1	0	0	0	1	1	7.9
PF13D12.6	CEH-8	1	1	1	1	0	0	0	1	1	7.9
PF13D12.6	DAF-3	1	1	1	1	0	0	0	1	1	7.9
PF13D12.6	DIE-1	1	1	1	1	0	0	0	1	1	7.9
PF13D12.6	DSC-1	1	1	1	1	0	0	0	1	1	7.9
PF13D12.6	ELT-2	1	1	1	1	1	1	1	1	0	7.9
PF13D12.6	MLS-2	1	1	1	1	0	0	0	1	1	7.9
PF13D12.6	ODR-7	1	1	1	1	0	0	0	1	1	7.9
PF13D12.6	UNC-42	1	1	1	1	0	0	0	1	1	7.9
PF13D12.6	ZK867.1	1	1	1	1	0	0	0	1	1	7.9
PF55B11.4	ALR-1	1	1	1	1	0	0	0	1	1	7.9
PF55B11.4	CEH-17	1	1	1	1	0	0	0	1	1	7.9
PF55B11.4	CEH-8	1	1	1	1	0	0	0	1	1	7.9
PF55B11.4	MLS-2	1	1	1	1	0	0	0	1	1	7.9
PF55B11.4	UNC-42	1	1	1	1	0	0	0	1	1	7.9
PF59F5.2	CEH-20	1	1	1	1	0	0	0	1	1	7.9
PF59F5.2	NHR-6	1	1	1	1	0	0	0	1	1	7.9
PF59F5.2	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pfar-7	CEH-8	1	1	1	1	0	0	0	1	1	7.9
Pfar-7	ELT-1	1	1	1	1	0	0	0	1	1	7.9
Pfar-7	UNC-42	1	1	1	1	0	0	0	1	1	7.9
Pfar-7	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pfat-2	MAB-5	1	1	1	1	0	0	0	1	1	7.9
Pfat-2	ODR-7	1	1	1	1	0	0	0	1	1	7.9
Pfat-2	PHP-3	1	1	1	1	0	0	0	1	1	7.9
Pfat-2	TBX-8	1	1	1	1	0	0	0	1	1	7.9

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pfat-2	UNC-42	1	1	1	1	0	0	0	1	1	7.9
Pfat-2	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pfat-4	NHR-10	1	1	1	1	0	0	0	1	1	7.9
Pfat-4	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pfat-6	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pgpd-3	CEH-17	1	1	1	1	0	0	0	1	1	7.9
Pgpd-3	K08B12.2	1	1	1	1	0	0	0	1	1	7.9
Pgpd-3	ODR-7	1	1	1	1	0	0	0	1	1	7.9
PH32C10.3	NHR-6	1	1	1	1	0	0	0	1	1	7.9
PH32C10.3	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Phis-14	DAF-3	1	1	1	1	0	0	0	1	1	7.9
Phlh-15	PAL-1	1	1	1	1	1	1	1	0	1	7.9
Plbp-8	ALR-1	1	1	1	1	0	0	0	1	1	7.9
Plbp-8	CEH-17	1	1	1	1	0	0	0	1	1	7.9
Plbp-8	CEH-6	1	1	1	1	0	0	0	1	1	7.9
Plbp-8	CEH-8	1	1	1	1	0	0	0	1	1	7.9
Plpd-2	DAF-3	1	1	1	1	0	0	0	1	1	7.9
PM01B12.5	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-137	DIE-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-137	NHR-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-137	NHR-67	1	1	1	1	0	0	0	1	1	7.9
Pnhr-137	TBX-8	1	1	1	1	0	0	0	1	1	7.9
Pnhr-137	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-140	DIE-1	1	1	1	1	1	1	1	1	0	7.9
Pnhr-178	CEH-20	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	CEH-8	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	ELT-6	1	1	1	1	0	0	0	1	1	7.9

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pnhr-178	MAB-5	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	NHR-12	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	NHR-142	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	NHR-178	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	NHR-273	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	NHR-28	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	NHR-79	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	PAG-3	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	PHP-3	1	1	1	1	0	0	0	1	1	7.9
Pnhr-178	Y53C12C.1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-242	CEH-17	1	1	1	1	0	0	0	1	1	7.9
Pnhr-242	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-25	ALR-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-25	PEB-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-25	TBX-8	1	1	1	1	0	0	0	1	1	7.9
Pnhr-25	ZK867.1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-25	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-68	CEH-10	1	1	1	1	0	0	0	1	1	7.9
Pnhr-68	CEH-17	1	1	1	1	0	0	0	1	1	7.9
Pnhr-68	MLS-2	1	1	1	1	0	0	0	1	1	7.9
Pnhr-68	NHR-234	1	1	1	1	0	0	0	1	1	7.9
Pnhr-68	PHP-3	1	1	1	1	0	0	0	1	1	7.9
Pnhr-68	UNC-42	1	1	1	1	0	0	0	1	1	7.9
Pnhr-68	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-8	ELT-2	1	1	1	1	0	0	0	1	1	7.9
Pnhr-8	NHR-102	1	1	1	1	0	0	0	1	1	7.9
Pnhr-8	NHR-34	1	1	1	1	0	0	0	1	1	7.9

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pnhr-8	PAG-3	1	1	1	1	0	0	0	1	1	7.9
Pnhr-8	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	CEH-37	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	DAF-3	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	DIE-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	NHR-109	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	NHR-111	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	NHR-12	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	NHR-234	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	NHR-273	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	NHR-45	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	NHR-79	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	ODR-7	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	TTX-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
Pnhr-86	ZTF-8	1	1	1	1	0	0	0	1	1	7.9
Pref-2	NHR-86	1	1	1	1	1	1	1	0	1	7.9
Pref-2	ODR-7	1	1	1	1	1	1	1	0	1	7.9
Pstr-47	ODR-7	1	1	1	1	0	0	0	1	1	7.9
Pstr-47	ZTF-1	1	1	1	1	0	0	0	1	1	7.9
PT09F3.1	T09F3.1	1	1	1	1	0	0	0	1	1	7.9
Ptag-257	PHP-3	1	1	1	1	1	1	1	0	1	7.9
PY57A10A.27	CEH-48	1	1	1	1	0	0	0	1	1	7.9
PY57A10A.27	CEH-8	1	1	1	1	0	0	0	1	1	7.9
PY57A10A.27	MAB-5	1	1	1	1	0	0	0	1	1	7.9
PY57A10A.27	PHP-3	1	1	1	1	0	0	0	1	1	7.9
PY57A10A.27	Y47D3B.9	1	1	1	1	0	0	0	1	1	7.9

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
PZK686.4	CEH-8	1	1	1	1	0	0	0	1	1	7.9
Phlh-30	FLH-1	0	1	1	1	1	1	1	1	N/A	8.3
Pptc-1	DIE-1	0	1	1	1	1	1	1	1	N/A	8.3
Psbp-1	ZTF-1	0	1	1	1	1	1	1	1	N/A	8.3
PT23F11.4	ZTF-1	0	1	1	1	1	1	1	1	N/A	8.3
PZK593.3	NHR-2	0	1	1	1	1	1	1	1	N/A	8.3
PZK593.3	NHR-91	0	1	1	1	1	1	1	1	N/A	8.3
PC09G9.7	TBX-8	1	1	1	1	0	1	0	1	1	8.6
PC36A4.9	ZC204.12	1	1	0	1	1	1	0	1	1	8.6
Pcyp-35A5	LIN-54	1	1	0	1	1	1	0	1	1	8.6
Pech-6	ZC204.12	1	1	0	1	1	1	0	1	1	8.6
Pegg-2	CEH-6	1	1	1	1	0	1	0	1	1	8.6
Pegg-2	ODR-7	1	1	1	1	0	1	0	1	1	8.6
Pegg-2	Y47D3B.9	1	1	1	1	0	1	0	1	1	8.6
PF13D12.6	C18G1.2	1	1	1	1	0	1	0	1	1	8.6
PF13D12.6	ELT-3	1	1	1	1	0	1	0	1	1	8.6
Pfat-6	LIN-54	1	1	0	1	1	1	0	1	1	8.6
PM01B12.5	LIN-54	1	1	0	1	1	1	0	1	1	8.6
PM01B12.5	Y53G8AM.8	1	1	0	1	1	1	0	1	1	8.6
PM01B12.5	ZIP-4	1	1	1	1	0	1	0	1	1	8.6
Pnhr-137	LIN-54	1	1	0	1	1	1	0	1	1	8.6
Pnhr-178	NHR-102	1	1	1	1	0	1	0	1	1	8.6
Pnhr-178	NHR-45	1	1	1	1	0	1	0	1	1	8.6
Pnhr-178	NHR-67	1	1	1	1	0	1	0	1	1	8.6
Pnhr-25	CEH-26	1	1	1	1	0	1	0	1	1	8.6
Pnhr-25	Y38C9A.1	1	1	0	1	1	1	0	1	1	8.6
Pnhr-86	NHR-28	1	1	1	1	0	1	0	1	1	8.6

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pnhr-86	NHR-66	1	1	1	1	0	1	0	1	1	8.6
Pnhr-86	PEB-1	1	1	1	1	0	1	0	1	1	8.6
PC06E7.3	CEH-20	1	1	1	1	1	1	0	1	1	9.3
PC30F12.1	CES-1	1	1	1	1	1	1	0	1	1	9.3
PC49C3.3	CEH-10	1	1	1	1	1	1	0	1	1	9.3
Pcyp-35A3	DIE-1	1	1	1	1	1	1	0	1	1	9.3
Pdhs-25	DAF-3	1	1	1	1	1	1	0	1	1	9.3
Pdop-3	ZTF-1	1	1	1	1	1	1	0	1	1	9.3
Pech-6	ZTF-1	1	1	1	1	1	1	0	1	1	9.3
Pegg-2	LIN-54	1	1	0	1	1	1	1	1	1	9.3
Pegg-2	ZTF-1	1	1	1	1	1	1	0	1	1	9.3
Pelo-2	CEH-6	1	1	1	1	1	1	0	1	1	9.3
Pepn-1	LIR-1	1	1	1	1	1	1	0	1	1	9.3
PF13D11.1	ODR-7	1	1	1	1	1	1	0	1	1	9.3
PF13D12.6	EGL-18	1	1	1	1	1	1	0	1	1	9.3
PF13D12.6	ELT-1	1	1	1	1	1	1	0	1	1	9.3
PF13D12.6	ELT-6	1	1	1	1	1	1	0	1	1	9.3
PF59F5.2	ODR-7	1	1	1	1	1	1	0	1	1	9.3
Pfat-2	DIE-1	1	1	1	1	1	1	0	1	1	9.3
Pgpd-3	ZTF-3	1	1	1	1	1	1	0	1	1	9.3
Phlh-30	ZTF-1	0	1	1	1	1	1	1	1	1	9.3
Plbp-8	MAB-5	1	1	1	1	1	1	0	1	1	9.3
Plbp-8	UNC-42	1	1	1	1	1	1	0	1	1	9.3
PM01B12.5	Y47D3B.9	1	1	1	1	1	1	0	1	1	9.3
Pnhr-178	ODR-7	1	1	1	1	1	1	0	1	1	9.3
Pnhr-25	F08G12.3	1	1	0	1	1	1	1	1	1	9.3
Pnhr-68	LIN-54	1	1	0	1	1	1	1	1	1	9.3

Promoter (bait)	Interacting TF (prey)	Self-active *lacZ*	Self-active *HIS3*	wTF2.1	Sticky *HIS3/lacZ*	Found multiple times	Found once from AD-TF mini library	Found with both libraries	Retest *HIS3*	Retest *lacZ*	Normalized score
Pnhr-68	ODR-7	1	1	1	1	1	1	0	1	1	9.3
Pnhr-86	NHR-70	1	1	1	1	1	1	0	1	1	9.3
Pnhr-86	NHR-86	1	1	1	1	1	1	0	1	1	9.3
Psrj-34	NHR-86	1	1	1	1	1	1	0	1	1	9.3
PY57A10A.27	ZTF-1	1	1	1	1	1	1	0	1	1	9.3
PZK686.4	NHR-6	1	1	1	1	1	1	0	1	1	9.3
Pacdh-2	ELT-1	1	1	1	1	1	1	1	1	1	10.0
Pacdh-2	NHR-10	1	1	1	1	1	1	1	1	1	10.0
Pech-6	DIE-1	1	1	1	1	1	1	1	1	1	10.0
Pegg-2	DIE-1	1	1	1	1	1	1	1	1	1	10.0
Pepn-1	ZTF-3	1	1	1	1	1	1	1	1	1	10.0
PF55B11.4	ELT-2	1	1	1	1	1	1	1	1	1	10.0
PF55B11.4	PHP-3	1	1	1	1	1	1	1	1	1	10.0
PF55B11.4	ZTF-1	1	1	1	1	1	1	1	1	1	10.0
Pfat-6	DIE-1	1	1	1	1	1	1	1	1	1	10.0
Pgpd-3	ZTF-1	1	1	1	1	1	1	1	1	1	10.0
PM01B12.5	DIE-1	1	1	1	1	1	1	1	1	1	10.0
Pmdt-15	ZTF-1	1	1	1	1	1	1	1	1	1	10.0
Pnhr-178	NHR-86	1	1	1	1	1	1	1	1	1	10.0
Pnhr-25	LSY-2	1	1	1	1	1	1	1	1	1	10.0
Pnhr-68	DIE-1	1	1	1	1	1	1	1	1	1	10.0
Psrj-34	HLH-30	1	1	1	1	1	1	1	1	1	10.0
PY57A10A.27	DIE-1	1	1	1	1	1	1	1	1	1	10.0

Table 2.4 List of interacting TFs identified in Y1H assays that occur in the modules

Module I	*Module II*	*Module III*	*Module IV*	*Module V*
T20F10.2	SEX-1	PAG-3	EGL-44	PHP-3
ZK867.1	NHR-10	NHR-79	ZTF-3	ZTF-8
LSY-2	NHR-41	Y53C12C.1	Y17G7B.20	ZTF-6
F26H9.2	NHR-2	NHR-273	VAB-3	TAB-1
MIG-5	NHR-1	NHR-12	NHR-43	NHR-6
F08G12.3	NHR-91	NHR-86	FLH-1	T13C5.4
PEB-1	PAL-1	NHR-28	PIE-1	ZTF-2
ALR-1	CES-1	NHR-45	T09F3.1	Y47D3B.9
	H02I12.5	NHR-234	HLH-30	ODR-7
		MAB-5	LIR-1	TBX-8
		NHR-111	K08B12.2	Y38C9A.1
		NHR-67	CEH-48	ZC204.12
		NHR-102	DAF-3	ZTF-1
		ELT-6	FLH-2	UNC-42
		CEH-20	ELT-3	CEH-10
			ELT-1	LIN-54
			ELT-2	C32D5.1
				DIE-1
				CEH-6
				CEH-8
				CEH-17
				MLS-2
				DSC-1

Preface to Chapter III

In this chapter, we describe the mapping of a protein-protein interaction network that consists of a specific cofactor and *C. elegans* TFs. We also describe the relevance of this network to *C. elegans* lipid metabolism and energy homeostasis. The work presented here includes contributions from Colin C. Conine, who generated the transgenic nematode strains used in this study.

Most of this chapter has been published separately in the following manuscript:

Arda, H. E., Taubert, S., MacNeil, L. T., Conine, C. C., Tsuda, B., Van Gilst, M., Sequerra, R., Doucette-Stamm, L., Yamamoto, K. R., and Walhout, A. J. M. (2010). Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. Mol. Syst. Biol *6*, 367.

CHAPTER III

A Cofactor-TF Network that Relates to *C. elegans* Lipid Metabolism

Abstract

Upon binding to their genomic sites, NHRs nucleate the assembly of multifactor transcriptional regulatory complexes by recruiting gene- and cell-specific cofactors. Studies using vertebrate models discovered several cofactor-NHR interactions that are crucial for energy homeostasis and physiology. Despite the existence of numerous NHRs in *C. elegans*, only a few *C. elegans* cofactor-NHR interactions have been identified. In this chapter, we investigate the protein-protein interactions involving the Mediator subunit MDT-15 and *C. elegans* TFs using high-throughput yeast two-hybrid (Y2H) assays, and analyze how this 'mini' cofactor-TF network relates to *C. elegans* metabolism by testing some of the interactions *in vivo*.

Introduction

In mammalian systems, transcriptional cofactor protein complexes, such as histone modifiers, chromatin remodelers and proteins that interact with the basal transcriptional machinery, can function as 'molecular adaptors' of physiologic pathways. Cofactors interact with several NHRs in a context dependent manner to mediate intra- and extra-cellular signals and modulate the transcriptional activity of NHRs on their target genes (Rosenfeld et al., 2006). For instance, by establishing physical contacts with mostly PPARs, *PPAR coactivator 1* (PGC-1) family of coactivators regulate mitochondrial biogenesis, and differentiation of brown fat tissue (Lin et al., 2005). Similarly, interactions involving members of the Mediator complex and mammalian TFs, such as MED1/PPARγ and MED15/SREBP-1, regulate adipogenesis, cholesterol and lipid homeostasis (Ge et al., 2002; Yang et al., 2006).

Thus far, only two *C. elegans* transcriptional cofactors have been characterized in detail; DIN-1, and MDT-15 (Ludewig et al., 2004; Taubert et al., 2006; Yang et al., 2006). MDT-15, the MED15 ortholog, interacts with NHR-49 and the SREBP-1 ortholog, SBP-1 (Ludewig et al., 2004; Taubert et al., 2006; Yang et al., 2006). Expression profiling of *mdt-15(RNAi)* and *nhr-49(RNAi)* animals showed that there is not a complete overlap between the targets of the two factors. MDT-15 is required for the expression of several additional metabolic

genes that are not targets of NHR-49. This implicates that MDT-15 interacts with other TFs, possibly to regulate distinct sets of genes.

In light of our findings in Chapter II, we hypothesized that MDT-15 may interact with the NHRs that we identified in the metabolic GRN. We used high-throughput Y2H matrix experiments to survey the ability of *C. elegans* TFs to interact with MDT-15. Below, we present the results of these experiments and show their relevance *in vivo*.

Results

MDT-15 specifically interacts with NHRs that occur in the metabolic GRN

The observation that expression of several metabolic genes requires MDT-15 function but not NHR-49 led to the hypothesis that MDT-15 likely interacts with additional TFs to exert its metabolic functions. We performed comprehensive Y2H assays using MDT-15 as bait *versus* our array of 755 full-length *C. elegans* TFs (80% of all 940, Figure 3.1A) (Reece-Hoyes et al., 2005; Vermeirssen et al., 2007). We did not have a functional clone for SBP-1, a known MDT-15 partner, (data not shown), but did confirm the interaction between MDT-15 and NHR-49. Additionally, we identified 12 novel interactions between MDT-15 and *C. elegans* TFs (Figure 3.1B and Table 3.1).

MDT-15 plays a central role in systems physiology. Therefore, we predicted that the TFs it interacts with should occur in our network. Indeed, statistical analysis of the TFs that interact with MDT-15 revealed three distinct features: they are significantly enriched for NHRs, for TFs that confer a Nile Red staining phenotype by RNAi, and for TFs that occur in the metabolic GRN (Figure 3.2 and Table 3.2). This confirms the prediction that MDT-15 interacts with many TFs and provides support for its central role in metabolic gene expression networks.

NHR-10 interacts with MDT-15 to regulate the expression of acdh *genes*

The expression of the *C. elegans* short-chain acyl-CoA dehydrogenases *acdh-1* and *acdh-2* requires MDT-15, but not its partner NHR-49 (Taubert et al., 2006). In the metabolic GRN that we described in Chapter II (Table 2.3), we identified NHR-10 as the single NHR that can bind to the promoters of both *acdh-1* and *acdh-2* (*Pacdh-1* and *Pacdh-2*). In addition, NHR-10 can interact with MDT-15 (Table 3.1). To determine if NHR-10 regulates *acdh-1* and/or *acdh-2 in vivo*, we created transgenic animals expressing GFP under the control of *Pacdh-1* or *Pacdh-2.* We noticed that the promoter of *acdh-1* is only active in the intestine, whereas the promoter of *acdh-2* is active in the intestine, hypodermis and body wall muscle. We then used RNAi to knock down the activity of either *mdt-15* or *nhr-10*. We observed that RNAi of *mdt-15* greatly reduced GFP expression in the intestine of both *Pacdh-1::GFP* and *Pacdh-2::GFP* animals, confirming that these promoters are MDT-15 targets (Taubert et al., 2006) (Figures 3.3A and B). Importantly, knockdown of *nhr-10* conferred an identical *acdh* expression phenotype strongly suggesting that NHR-10 and MDT-15 together activate *acdh-1* and *acdh-2* expression by binding to their promoters *in vivo* (Figures 3.3A and B). Altogether, these data show that MDT-15 preferentially associates with NHRs that occur in the metabolic GRN to exert its functions.

Discussion

In this chapter, we showed that MDT-15, which was previously identified as a metabolic *C. elegans* transcriptional cofactor, preferentially associates with NHRs, and specifically with those that occur in the metabolic GRN.

The observation that MDT-15 partners with only 13 TFs out of 755 suggests that the protein-protein interactions that we have identified are specific and the network is of high quality. However, it is possible that due to context dependent post-translational modifications of TFs, we might have missed additional interactions, such as the previously reported interaction between MDT-15 and SBP-1. Nevertheless, our assays successfully identified a novel partnership between MDT-15 and NHR-10, which regulates a distinct set of metabolic genes than the MDT-15/NHR-49 complex. Our findings support the initial hypothesis that MDT-15 acts as a central coactivator subunit that forms complexes with a diverse set of TFs to regulate various aspects of *C. elegans* metabolism.

Homologs of PGC-1 family are found in vertebrates but not in yeast, flies or nematodes (Lin et al., 2005). Invertebrates also lack the conserved LXXLL motif that is found in vertebrate Mediator subunits; instead, most of the cofactor-TF interactions are mediated by the 'KIX domain' (Näär and Thakur, 2009). Interestingly, a recent study describes a new interaction between PGC-1α and the murine TF, TWIST-1 that controls brown fat metabolism (Pan et al., 2009). In

our Y2H analysis, we found HLH-8, the *C. elegans* ortholog of Twist-1, as one of the MDT-15 interacting TFs. This suggests that although the domains necessary for these interactions are different, the interaction between a metabolic cofactor and the orthologous TF is conserved.

In the future, it will be necessary to extend systematic protein-protein interaction mapping between *C. elegans* cofactors and TFs to genome-scale to further our understanding of cofactor-TF regulation in the nematode systems physiology.

Materials and Methods

Y2H matrix assays

Y2H matrix assays were performed by mating as described (Walhout and Vidal, 2001; Vermeirssen et al., 2007). The DB-MDT-15-NT construct (Taubert et al., 2006) was transformed into MaV103 yeast and used as bait. This bait strain was mated against the *C. elegans* TF array (wTF2.1), which consists of 755 TFs (Vermeirssen et al., 2007). Diploids were selected on permissive media (Sc-Leu-Trp) and subsequently replica plated onto selective media (Sc-His-Leu-Trp + 20mM 3-amino-1,2,4-triazole) to assay for *HIS3* reporter gene activity. Diploids that grew on these plates were picked and assayed for *LacZ* reporter gene activation. Only TFs that conferred positive read-outs in both reporter gene assays were considered as interacting TFs. The identities of the interacting prey TFs were determined based on their position in the array, and confirmed by sequencing.

Generation of promoter::GFP transgenic C. elegans and quantification of GFP expression

Transgenic animals were created by ballistic transformation into *unc-119(ed3)* animals as described (Reece-Hoyes et al., 2007). The promoter sequences used to generate *promoter::GFP* constructs were identical to those used in Y1H assays. For each GFP construct we obtained up to ten independent

lines that all displayed identical GFP expression patterns (data not shown). When available an integrated, or alternatively the best transmitting line, was selected for subsequent experiments. GFP expression was scored for each genotype in at least 25 animals, and each experiment was repeated at least twice.

RNAi experiments

RNAi clones were cherry-picked from the *C. elegans* RNAi library (Kamath et al., 2003). HT115(DE3) bacterial strains carrying RNAi constructs were grown in Luria Broth containing ampicillin (50 µg/mL) at 37°C for 8-10 hours until they reach OD600 ~1.0. Bacteria were pelleted by centrifugation, washed once with M9 buffer, resuspended in M9 buffer, and seeded on NGM plates containing 5 mM IPTG and 50 µg/mL ampicillin. Bacteria were induced overnight at 22°C. The next day, four to five L4 larvae were placed on the plates. All RNAi clones were verified by sequencing.

Microscopy

Nomarski and fluorescence images were obtained using a Zeiss Axioscope 2+ microscope. Images of GFP expression were captured using a digital CCD camera (Hamamatsu C4742-95-12ERG) and Axiovision (Zeiss) software. GFP fluorescence images were obtained using a FITC filter (excitation 460-500 nm, emission 510-565 nm). Animals were placed into a drop of 0.1% sodium azide in M9 buffer on a fresh 2% agarose pad for observation.

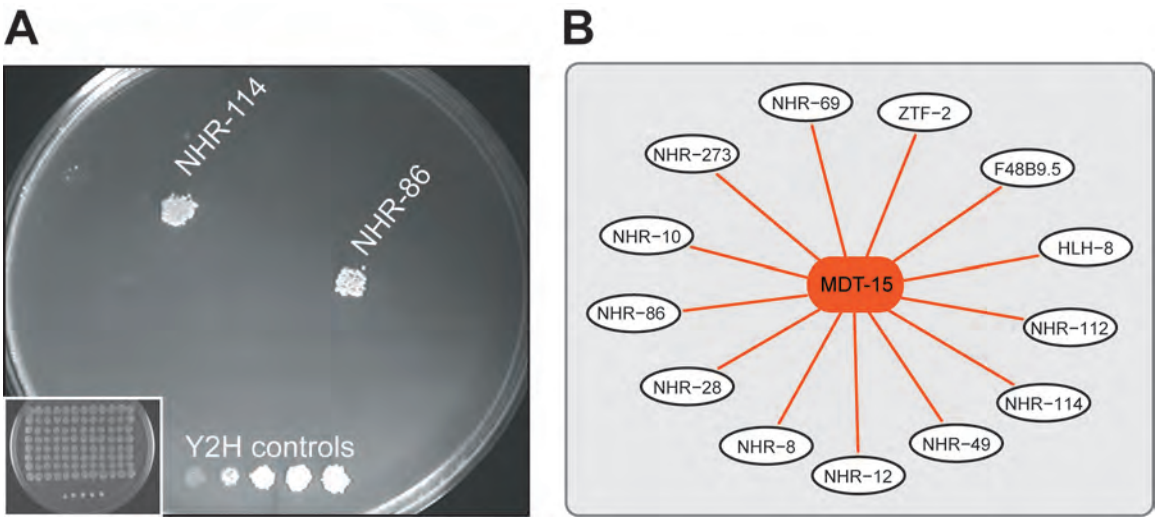

Figure 3.1 MDT-15 can interact with additional *C. elegans* TFs.

(A) Example of a Y2H matrix experiment using DB-MDT-15 as bait. Indicated are the TFs that interact with MDT-15, which can grow on the selective plate, inset - permissive plate. Five spots at the bottom are Y2H controls.

(B) The protein-protein interaction map of MDT-15 and *C. elegans* TFs. Each white node is a *C. elegans* TF, and orange lines represent mapped Y2H interactions.

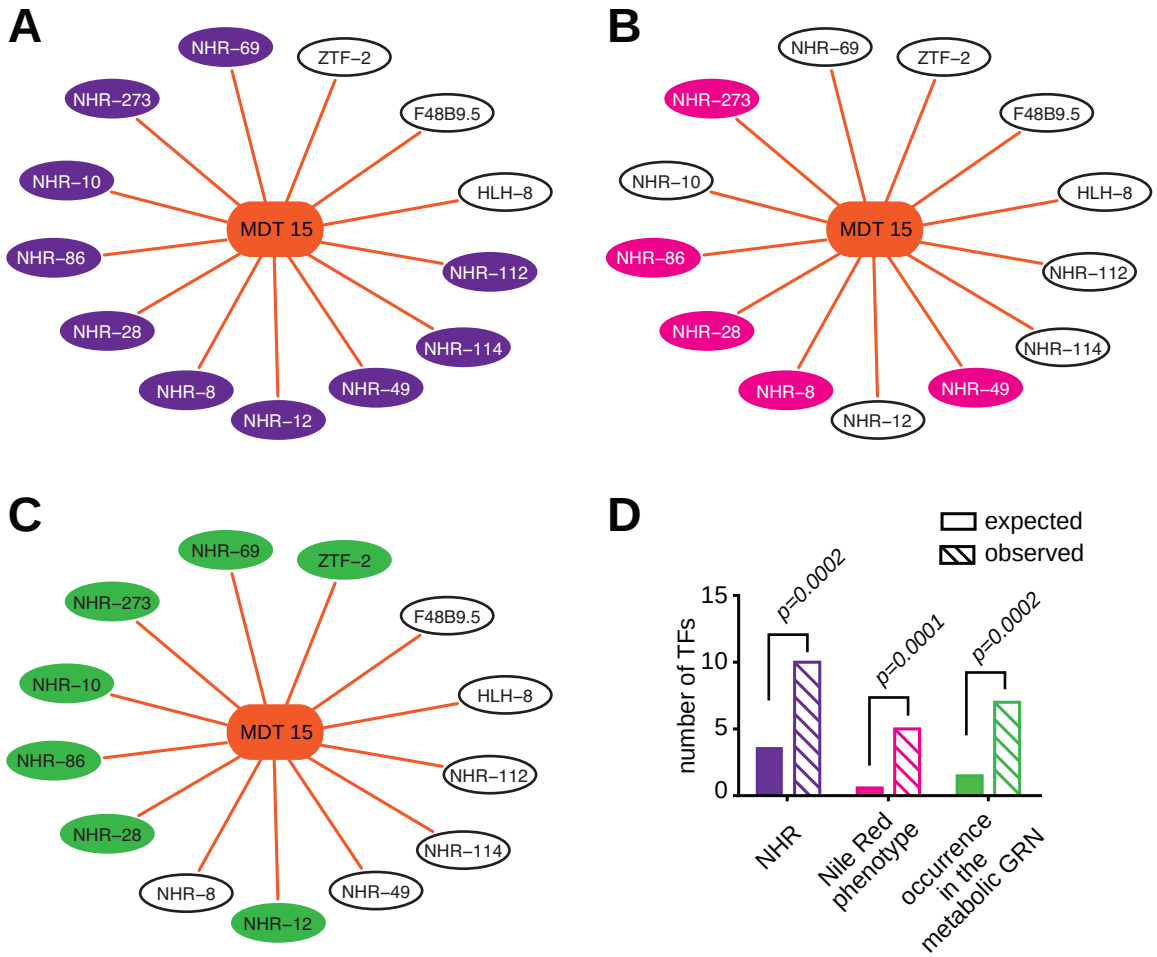

Figure 3.2 Common characteristics of the MDT-15 interacting TFs.

Protein-protein interaction maps of MDT-15 and *C. elegans* TFs are color coded to indicate different attributes.

(A) TFs that have NHR type DNA binding domain - purple.

(B) TFs that confer Nile Red staining phenotype - magenta.

(C) TFs that occur in the metabolic GRN – green.

(D) Bar graphs show the expected (solid bars) and observed (dashed bars) frequencies of MDT-15 interacting TFs that have indicated attributes.

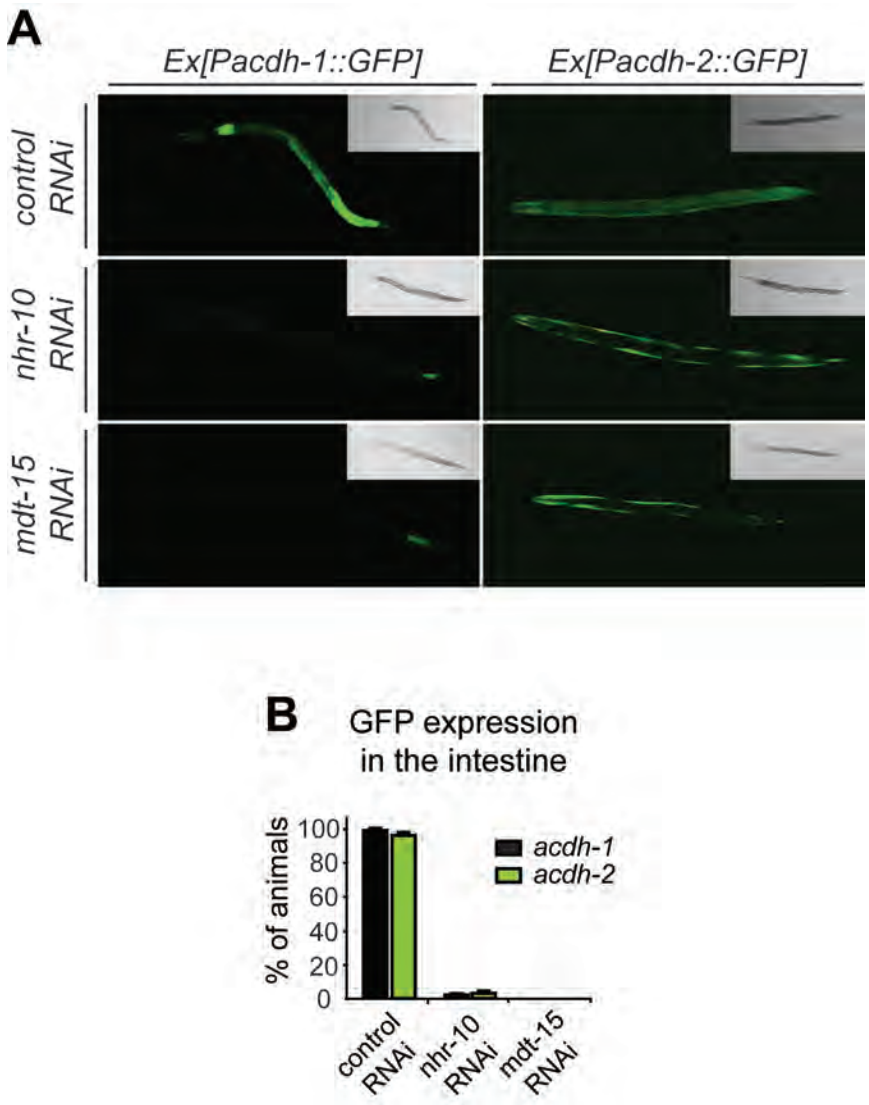

Figure 3.3 MDT-15 and NHR-10 regulate the activity of *Pacdh-1* and *Pacdh-2.*

(A) Images showing GFP expression in *Pacdh-1::GFP* (left) and *Pacdh-2::GFP* (right) transgenic animals that were subjected to control *RNAi* (top), *nhr-10 RNAi* (middle), and *mdt-15 RNAi* (bottom). Remaining GFP expression in *Pacdh-2::GFP* animals occurs in body wall muscle. Insets in the top right corners correspond to the DIC image of the animal.

(B) Bar graphs representing quantifications of intestinal GFP expression shown in (A). Error bars indicate the SEM.

Table 3.1 List of TFs that interact with MDT-15 in Y2H assays

Bait	TF Prey	DNA-binding domain	Confers a Nile Red phenotype?		In the metabolic GRN as an interacting TF?
DB-MDT-15-NT	ZTF-2	ZF - C2H2	not tested		yes
	NHR-10	ZF-NHR	no		yes
	NHR-69	ZF-NHR	not tested		yes
	NHR-12	ZF-NHR	no		yes
	F48B9.5	Paired Domain - NPAX	not tested		no
	NHR-114	ZF-NHR	not tested		no
	NHR-86	ZF-NHR	yes [1]	(increased)	yes
	NHR-28	ZF-NHR	yes [1]	(increased)	yes
	NHR-273	ZF-NHR	yes [1]	(increased)	yes
	NHR-49	ZF-NHR	yes [2]	(increased)	no
	NHR-8	ZF-NHR	yes [2]	(increased)	no
	NHR-112	ZF-NHR	not tested		no
	HLH-8	bHLH	not tested		no

[1] *this study*
[2] *Ashrafi et al., 2003.*

Table 3.2 Contingency table analysis of TFs that interact with MDT-15

NHR association

Binds MDT-15	NHRs in wTF2.1 array	non-NHRs in wTF2.1 array
Yes	10	3
No	195	547

(Fisher's exact, two-tailed p=0.0002)

Nile Red phenotype association

Binds MDT-15	TFs with Nile Red phenotypes in wTF2.1 array	TFs without Nile Red phenotypes in wTF2.1 array
Yes	5	8
No	29	713

(Fisher's exact, two-tailed p=0.0001)

Metabolic GRN association

Binds MDT-15	In the metabolic GRN	Not in the metabolic GRN
Yes	7	6
No	80	662

(Fisher's exact, two-tailed p=0.0002)

Preface to Chapter IV

This chapter describes an integrated NHR network that regulates lipid metabolism of *C. elegans* as well as the *in vivo* analysis of protein-DNA interactions involving the NHRs in the integrated network. The work presented here includes contributions from Dr. Lesley T. MacNeil, who performed some of the genetic rescue experiments.

Most of this chapter has been published separately in the following manuscript:

Arda, H. E., Taubert, S., MacNeil, L. T., Conine, C. C., Tsuda, B., Van Gilst, M., Sequerra, R., Doucette-Stamm, L., Yamamoto, K. R., and Walhout, A. J. M. (2010). Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. Mol. Syst. Biol *6*, 367.

CHAPTER IV

An Integrated NHR Network that Regulates *C. elegans* Lipid metabolism

Abstract

Delineating physical protein-DNA or protein-protein interactions are essential for generating GRN models. However, current technologies to map these interactions tend to yield static networks that do not immediately reflect the dynamic state of the system. Therefore, it is important to generate ‘integrated networks’ by combining physical interaction maps with other types of data, such as expression and phenotype analysis. In this chapter we have generated an integrated NHR network using the protein-DNA and the protein-protein interaction data, as well as the lipid staining phenotypes of the NHRs that we have described in previous chapters. Using the integrated NHR network, we made several regulatory and functional predictions, which we further tested *in vivo.*

Introduction

Arguably, one of the most important aspects of network biology studies is to be able to predict functional relationships between the components of a mapped network that can occur in biological context [reviewed in (Piano, 2006)]. To date, several types of networks relating to *C. elegans* biology have been mapped. These include but are not limited to, networks of strictly physical protein-DNA or protein-protein interactions, and phenotypic interactions (Deplancke et al., 2006; Vermeirssen et al., 2007a; Martinez et al., 2008; Grove et al., 2009; Li et al., 2004; Gunsalus et al., 2005). In order to find global correlations between different functional characteristics of a biological system, integrated networks are constructed in which different data types are incorporated into a single graph model [reviewed in (Lehner and Lee, 2008)]. Using this model, researchers can predict novel gene functions and phenotypes.

Another related challenge in the network field is to translate mapped physical interactions into regulatory ‘edges’. The physical mapping methods that we have used reveal which interactions can happen, but they do not directly reveal when and where these interactions can occur in a biological setting. Thus, it is important to extend our efforts to investigate the regulatory consequences of mapped interactions *in vivo*, to further our understanding of the dynamic state of biological networks.

Motivated by these arguments, we generated an integrated NHR network that contains protein-DNA interactions, protein-protein interactions, as well as metabolic phenotypes relating to the NHRs that we have described in previous chapters. This network allowed us to visualize the relationships of these different parameters to each other, and make predictions regarding *C. elegans* energy homeostasis, which we then tested *in vivo*.

Results

An integrated NHR network

We visualized all NHR-promoter interactions that contribute to modules II and III, the protein-protein interactions involving MDT-15, as well as the metabolic phenotypes in a single, integrated NHR network (Figure 4.1). We also included previously identified protein-protein interactions between NHR-49 and any NHR that is present in the network (Vermeirssen et al., 2007b). Several observations can be made from this GRN. First, MDT-15 interacts with NHRs from both module II and III, and also interacts with TFs that receive inputs from either module. This further emphasizes the central role of MDT-15 in the regulation of metabolic gene expression. Second, three NHRs in module III can physically interact with NHR-49 (Vermeirssen et al., 2007b), and RNAi of all of them results in a fat storage phenotype, suggesting that NHR-49 shares different partners for its metabolic functions. Third, there are numerous interactions between NHRs and NHR-encoding genes, and several *nhr* gene promoters receive input from multiple other NHRs. For instance, *Pnhr-49* interacts with four NHRs, both from module II and III. This 'interregulation' implies that NHRs may function in transcriptional cascades to execute appropriate gene expression programs, *e.g.* upon receiving intrinsic or extrinsic signals (Magner and Antebi, 2008). Fourth, several NHRs interact with multiple promoters, which may ensure a coordinated response of these target genes. Fifth, there are multiple feed-

forward loops in the network, in which an NHR controls another NHR, and both of these share a downstream target. Such loops likely enable a controlled signal response, and protect against fluctuations in gene expression (Alon, 2007). Finally, NHR-178 and NHR-86 can interact with their own promoters, suggesting that they may be auto-regulators.

NHR-86: a metabolic TF with high flux-capacity

The HNF4 homolog NHR-86 (module III) has a high flux-capacity (Martinez et al., 2008): its promoter is bound by twelve NHRs and it interacts with seven gene promoters. In addition, NHR-86 interacts with MDT-15 and with its own promoter (Figure 4.1). In the *Saccharomyces cerevisiae* GRNs, nodes with high incoming degree, i.e. 'target hubs', are key regulators of differentiation pathways (Borneman et al., 2006). Thus, we hypothesized that NHR-86 might be a novel regulator of *C. elegans* metabolic pathways. We showed in Chapter II that inactivation of *nhr-86* results in increased Nile Red and Oil-Red-O staining. In order to further characterize *nhr-86*, we chose to use a deletion mutant, which we requested from the Japanese National Bioresource Consortium.

tm2590 *is a null allele of* nhr-86

The *tm2590* allele results from a 172bp deletion and insertion of the nucleotides "GG" at the 3' end of the fourth exon of the *nhr-86* gene (Figure 4.2A). We determined by cDNA cloning and sequencing that this deletion and

insertion results in the production of a longer transcript in *tm2590* animals due to abnormal splicing and retainment of the fourth intron. However, the inclusion of the fourth intron introduces a premature stop codon, which causes a truncation at the ligand binding domain of the NHR-86 protein (Figure 4.2B).

We wanted to test if this truncated form of NHR-86 can still bind to its DNA target in Y1H assays. Using cDNA synthesized from mutant animals, we generated an AD-ΔNHR-86 plasmid by Gateway cloning (Walhout et al., 2000; Walhout and Vidal, 2001). We tested the interaction by transforming the plasmids that have either wild-type or mutant NHR-86 into the *Pnhr-86* Y1H bait. We found that while wild-type NHR-86 can bind to its own promoter in Y1H assays, truncated NHR-86 cannot (Figure 4.2C), even though both forms of NHR-86 are stably expressed in yeast (Figure 4.2D). This may imply that the ligand binding domain of NHR-86 is necessary for its ability to bind its DNA target in yeast, or a more likely explanation is that the truncation impairs proper folding and consequently the DNA binding ability of NHR-86.

We also raised peptide antibodies against NHR-86, and assayed total protein extracts from wild-type and *nhr-86(tm2590)* animals. Despite the fact that the truncated form of NHR-86 is expressed in yeast, we did not detect any NHR-86-specific protein in the mutant by western blotting or immunoprecipitation assays, suggesting that the mutant protein might be subjected to surveillance programs (Figure 4.3A and B). Taken together, our results show that *tm2590*

animals cannot produce a wild-type NHR-86 protein and that *tm2590* is likely a null allele of *nhr-86.*

Is NHR-86 an autoregulator?

The observation that NHR-86 interacts with its own promoter in Y1H assays suggests that it may have auto-regulatory activity. To test this, we created transgenic animals expressing GFP under the control of *Pnhr-86* and examined GFP expression in wild-type and *nhr-86(tm2590)* mutant animals. Compared to wild-type *Pnhr-86::GFP* animals, we found that GFP expression was substantially upregulated in the pharynx and hypodermis in 100% of the *nhr-86(tm2590)* mutant (Figure 4.4A). This suggests that NHR-86 is an auto-repressor in these tissues. Negative auto-regulation is an important feature of regulatory circuits; it accelerates the response to outside signals and promotes transcriptional robustness (Alon, 2007).

Determining the expression pattern and subcellular localization of NHR-86

Determining the tissues or cellular compartments in which the genes are expressed often provides important clues to understanding their biological function (Reece-Hoyes et al., 2007; Hunt-Newbury et al., 2007). To further investigate the expression pattern of NHR-86, we tagged the full-length *nhr-86*

ORF with GFP and generated transgenic animals[1]. We found that NHR-86 is a nuclear protein, which is in agreement with its function as a TF. Predominantly, NHR-86 localizes to the nuclei of intestinal and excretory gland cells, as well as several head neurons, which are all important tissues for metabolism and homeostasis (Figure 4.4B).

nhr-86(tm2590) *animals have abnormal amount of body fat*

Next, we examined the body fat levels of *nhr-86(tm2590)* mutants by both Nile Red and Oil-Red-O assays. We found that *nhr-86(tm2590)* animals exhibit an increase in Nile Red, as well as Oil-Red-O staining, confirming the *nhr-86 RNAi* result that we previously presented in Chapter II (Figures 2.6 and 4.5A, 4.5B). We also found that the Oil-Red-O phenotype was rescued when we crossed the mutant animals to the nematodes carrying the wild-type *nhr-86* construct (Figure 4.5B). Altogether our results show that NHR-86 functions to regulate lipid storage and/or catabolism.

nhr-86(tm2590) *animals have normal mean lifespan*

Metabolic pathways are important determinants of lifespan across the animal kingdom (Narasimhan et al., 2009). In *C. elegans*, the insulin-like receptor mutant *daf-2,* has an abnormally high level of body fat and live two times longer than wild-type animals. On the other hand, *nhr-49* mutants, which also

[1] See Appendix for nematode strain information

accumulate more body fat than wild-type animals, have shorter lifespan. Considering that *nhr-86(tm2590)* mutants have higher body fat content, we wanted to measure the average lifespan of *nhr-86(tm2590)* animals. Interestingly, unlike *daf-2* or *nhr-49*, we found that loss of *nhr-86* function did not affect the mean lifespan (Figure 4.5C).

An NHR circuit that responds to nutrient availability

The expansion of NHRs, their modular wiring in the network and the observation that they all contribute to systems physiology suggest that these NHRs may respond to physiological or environmental cues such as nutrient availability. In the integrated network, we showed that *nhr-178* is part of several gene circuits that involve additional NHRs in module III (Figure 4.1), Also in Chapter II we showed that reduction of *nhr-178* by RNAi results in increased body fat in the nematodes. One of the NHRs that interact with the promoter of *nhr-178* is NHR-45 (Figure 4.1). We obtained an *nhr-45(tm1307)* mutant, which has a large deletion in the DNA binding domain of *nhr-45* (Figure 4.6A). We found that the mutant recapitulates the Oil-Red-O staining profile observed with *nhr-45 RNAi* that we also presented in Chapter II (Figures 2.6 and 4.6B).

In order to further investigate this NHR circuit, we created a transgenic *C. elegans* strain that expresses GFP under the control of *Pnhr-178*[2], and found that this promoter mainly drives expression in the pharynx and the first anterior

[2] See Appendix for nematode strain information

intestinal cells (Int1) (100±0% of the animals, Figure 4.7A). When we crossed the *Pnhr-178::GFP* strain into the *nhr-45(tm1307)* animals, loss of *nhr-45* greatly reduced GFP expression in the pharynx and eliminated it in the Int1 cells (100±0%, Figure 4.7A). Microinjection of an *nhr-45* rescuing construct restored the intestinal and pharyngeal GFP expression (80±1% of the animals), demonstrating that NHR-45 activates the promoter of *nhr-178* in these tissues (Figure 4.7B). Interestingly, however, under standard well-fed conditions, in a small but reproducible number of *nhr-45(tm1307)* mutant animals (3.5±0.7%), we observed additional *nhr-178* promoter activity in the hypodermis, suggesting that NHR-45 may repress *Pnhr-178::GFP* expression in the hypodermis (Figure 4.7C). This observation, together with NHRs ability to act as molecular sensors, and the hypodermis being a lipid storage tissue led us to hypothesize that *nhr-45/nhr-178* gene circuit might be responsive to the nutrient state of the nematode.

Indeed, when we challenged wild-type animals with food withdrawal, *Pnhr-178::GFP* expression was specifically upregulated in the hypodermis (in 84±1% of the animals, Figure 4.8A), suggesting that this alternative expression pattern is part of a dynamic transcriptional program that responds to starvation. Supporting our conclusion, when we provided starved animals with food, GFP expression in the hypodermis rapidly disappeared (Figure 4.8A). In the case of *nhr-45(tm1307)* mutants, the starvation response was similar to that of wild-type animals. However, when they were provided with food after starvation the mutants failed

to completely suppress hypodermal *Pnhr-178* activity (90±0.3% of the animals still exhibited hypodermal expression, Figure 4.8A), again indicating that NHR-45 function is necessary for proper repression of *Pnhr-178* in this tissue.

Altogether, these observations reveal a gene circuit that involves multiple NHRs which function to rapidly respond to nutrient availability. In this circuit, *Pnhr-178* activity is kept off in the hypodermis under well-fed conditions. When animals starve, NHR-45 may be repressed, thereby allowing other TFs, again potentially NHRs, to activate hypodermal *nhr-178* expression (Figure 4.8B).

Discussion

In this chapter, we presented an integrated NHR network through which we analyzed regulatory interactions involving several previously uncharacterized NHRs, and showed that they are biologically relevant in the maintenance of *C. elegans* lipid homeostasis.

NHR-86, a TF with high-flux capacity, is an auto-repressor and required for keeping normal levels of body fat in the nematode. We also found that it is expressed in metabolically important tissues, such as the intestine. In addition, we have generated specific polyclonal antibodies for NHR-86, which work efficiently in immunoblotting and immunoprecipitation assays. In the future, these antibodies can be used for further biochemical characterization of the NHR-86 protein, and for the identification of its *in vivo* DNA targets using TF-centered methods, such as ChIP.

The integrated NHR network model, combined with our phenotypic analysis of the NHRs in modules led us to predict and uncover a novel NHR circuit involving *nhr-45* and *nhr-178* that responds to nutrient availability. Strikingly, we found that NHR-45 has two distinct transcriptional roles on the *Pnhr-178* activity; an activator in the intestine, yet a repressor in the hypodermis. It is important to note that, both of these NHRs engage in several other connections with additional NHRs, which we have not investigated. Thus, one explanation for the low penetrance of the hypodermal GFP expression in *nhr-*

45(tm1307) animals under well-fed conditions could be that the repression of *Pnhr-178* in this tissue is controlled by multiple TFs. Losing only the *nhr-45* activity might be enough to stagger the circuit, but not enough to completely inactivate it. As more mutant strains become available, it will be interesting to investigate the role of other NHRs in this nutrient sensing circuit by combining genetics and more rigorous phenotype testing, and delineate the combinatorial transcriptional regulation governed by these NHRs.

Materials and Methods

Microscopy

Nomarski and fluorescence images were obtained using a Zeiss Axioscope 2+ microscope. Images of GFP expression or Nile Red fluorescence were captured using a digital CCD camera (Hamamatsu C4742-95-12ERG) and Axiovision (Zeiss) software. GFP fluorescence images were obtained using a FITC filter (excitation 460-500 nm, emission 510-565 nm). Nile Red fluorescence images were taken using a rhodamine filter (excitation 525-555 nm, emission 575-630 nm). Animals were placed into a drop of 0.1% sodium azide in M9 buffer on a fresh 2% agarose pad for observation. Oil-Red-O images were acquired using Zeiss AxioCam HRc color CCD camera.

Generation of promoter::GFP transgenic C. elegans and quantification of GFP expression

Transgenic animals were created by ballistic transformation into *unc-119(ed3)* animals as described (Reece-Hoyes et al., 2007), with the exception that *Pnhr-86::nhr-86*^{ORF}*::GFP* was cloned by multi-site Gateway LR reaction into pDEST-MB14 (Dupuy et al., 2004). The promoter sequences used to generate *promoter::GFP* constructs were identical to those used in Y1H assays. For each GFP construct we obtained up to ten independent lines that all displayed identical GFP expression patterns (data not shown). When available an

integrated, or alternatively the best transmitting line, was selected for subsequent experiments. GFP expression was scored for each genotype in at least 25 animals, and each experiment was repeated at least twice.

Genetic rescue experiments

The *nhr-86* rescue strain was generated by crossing *nhr-86(tm2590)* mutants to the VL505 strain, which carries the *Pnhr-86::nhr-86*^{ORF}*::GFP* construct. The *nhr-45* rescuing fragment was PCR amplified using the forward primer CTCTTCATTATGCATTTTTGTTC and the reverse primer TCACTGGAAACGTGAGAGTCA from *C. elegans* (N2) genomic DNA. This fragment consists of the genomic sequences 2kb upstream of the translational start site of *nhr-45*, *nhr-45* gene itself, and 487 bp downstream of the 3'end of *nhr-45*. The PCR band corresponding to the expected size was gel purified and sequence verified. Transgenic animals were generated by microinjecting 10 ng/µL of the purified PCR product along with 80 ng/µL of the coinjection marker plasmid pRF4, which carries the marker gene *rol-6(su1006)* into the germline of the VL739 strain (see Appendix for a list of *C. elegans* strains generated in this study) (Mello et al., 1991).

Antibodies

We raised a polyclonal anti-NHR-86(NT) antibody using an N-terminal peptide (SQFRPEKKEKSTCSIC) and an anti-NHR-86(IN) antibody using an

internal peptide (SVGDGYPPTTSAQSALC) in rabbits (AnaSpec, Inc., San Jose, CA). anti-NHR-86(NT) was used at a final concentration of 1 µg/mL in immunoblotting assays. For detecting Gal4AD-fusions from yeast extracts, an anti-Gal4AD antibody from Sigma (cat #: G9293) was used.

Immunoblotting

C. elegans total protein extract preparation: L4 stage animals were collected and washed with sterile water. Animals were homogenized by sonicating in lysis buffer (50mM Tris-HCl pH 7.5, 150mM NaCl, 1mM EDTA, 1% Triton-X, 1mM DTT, and Roche Complete protease inhibitors). The lysates were cleared by centrifugation and protein concentrations were determined using Bio-Rad *Dc* Protein Assay (cat. #: 500-0116). 100 µg of total protein extract was loaded in each lane. Yeast sample preparation: Yeast strains were grown in YEPD liquid medium at 30°C with a starting density of OD600 ~0.1 until they reach an OD600 ~0.6-0.8. For each sample, 1 mL of liquid culture was pelleted and washed with sterile water. The pellet was then resuspended in freshly made, ice-cold 150 mM NaOH solution. After incubating on ice for 15 min, the cells were pelleted, resuspended in 20 µL of 2X Laemmli buffer, boiled for 5 minutes and loaded in equal amounts onto the gel. In both cases, proteins were separated using NuPAGE 4-12% Bis-Tris gels (Invitrogen, cat #: NP0323), and transferred onto PVDF membranes. The membranes were incubated overnight at 4°C with

primary antibodies, followed by standard immunoblotting techniques (Maniatis et al., 1982).

Immunoprecipitation (IP) assays

Total protein extracts of the nematodes were prepared and quantified as described above. 100 µg of each extract was saved as input samples. IP assays were performed with protein extracts adjusted to a final concentration of 5mg/mL. Extracts were pre-cleared with Sepharose A/G beads (Santa Cruz) at 4ºC for 1 hour, and transferred to fresh tubes for antibody incubation. 2 µg of antibody was added to each IP sample, and incubated on a nutator at 4ºC overnight. The next day, antibodies were bound to Sepharose A/G beads by nutating at 4ºC for an hour. The antibody bound beads were collected by brief centrifugation, boiled for 5 minutes in 2X Laemmli buffer, and loaded onto a 10% SDS-PAGE gel (Maniatis et al., 1982). The electrophoresis procedure is performed as described above.

Lifespan assays

Lifespan assays were performed as described (Oh et al., 2005). For each strain, approximately 70 late L4 larvae of each strain per experiment (15-20 animals/plate) were examined on nematode growth medium (NGM) plates containing 100µg/ml FUDR (5-fluoro-2′-deoxyuridine), and that were seeded with OP50 *E. coli* bacteria. Every two days, animals were examined for their ability to

respond to tapping with a platinum wire pick. The animals were scored dead when no movement could be observed. Each assay was repeated four times, and the identities of the strains were concealed while scoring to prevent observer bias.

Starvation assays

The assays were done as described (Van Gilst et al., 2005a, 2005b). Briefly, animals were collected and washed several times with M9 buffer to remove residual bacteria. The animals were then plated onto NGM plates that did not contain food, and allowed to fast on these plates for 12 hours at 20°C, and then scored for hypodermal GFP expression. After the starvation period, animals were transferred back on NGM plates with food and allowed to recover for 24 hours at 20°C, and then scored for hypodermal GFP expression. The data reported in Figure 4.8 was obtained by using synchronized L1 stage animals for starvation assays. Similar results were obtained with L3 stage animals (data not shown).

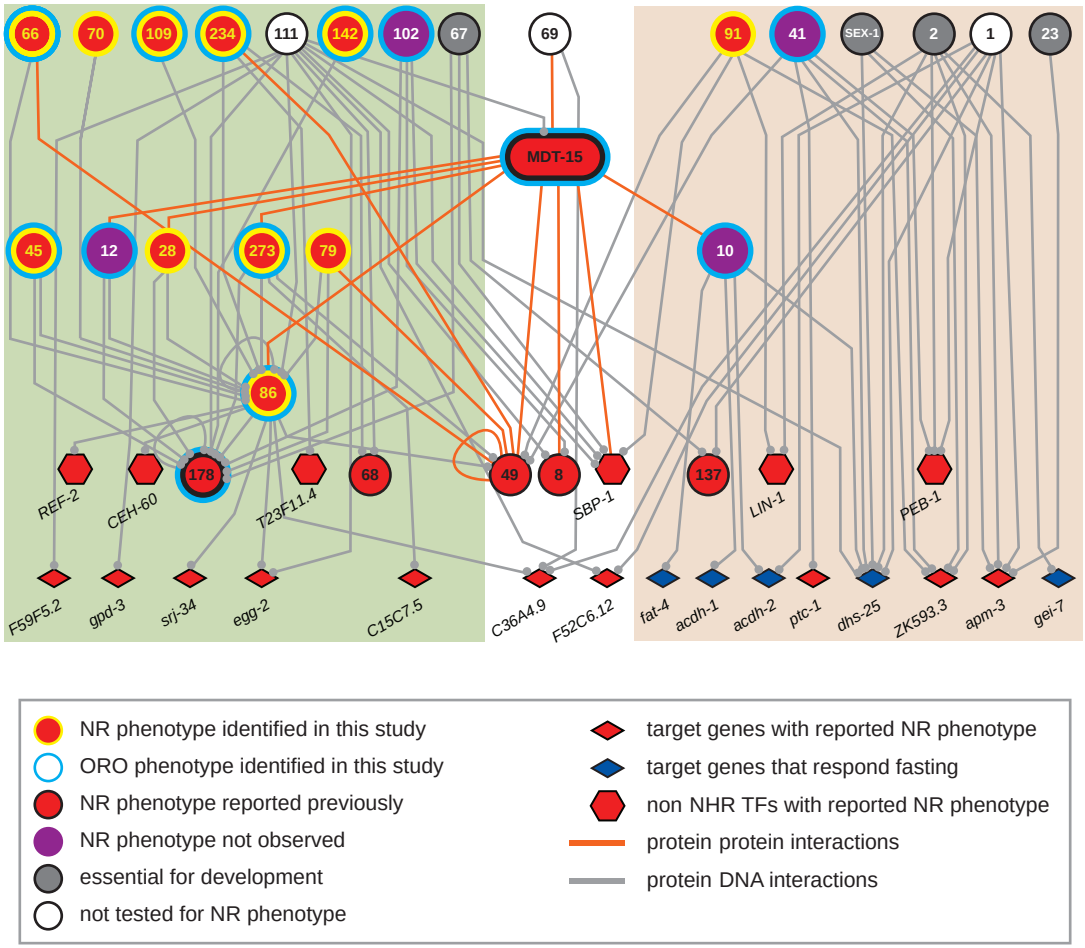

Figure 4.1 An integrated NHR network.

Interactions involving NHRs that are in module II (right, beige background) and module III (left, green background), their targets and MDT-15 are represented. NHRs that specifically bind a single target gene that is part of either module were also included. Circles - NHRs; numbers depict NHR identity, i.e. '1' is NHR-1; NR - Nile Red staining, ORO - Oil-Red-O staining.

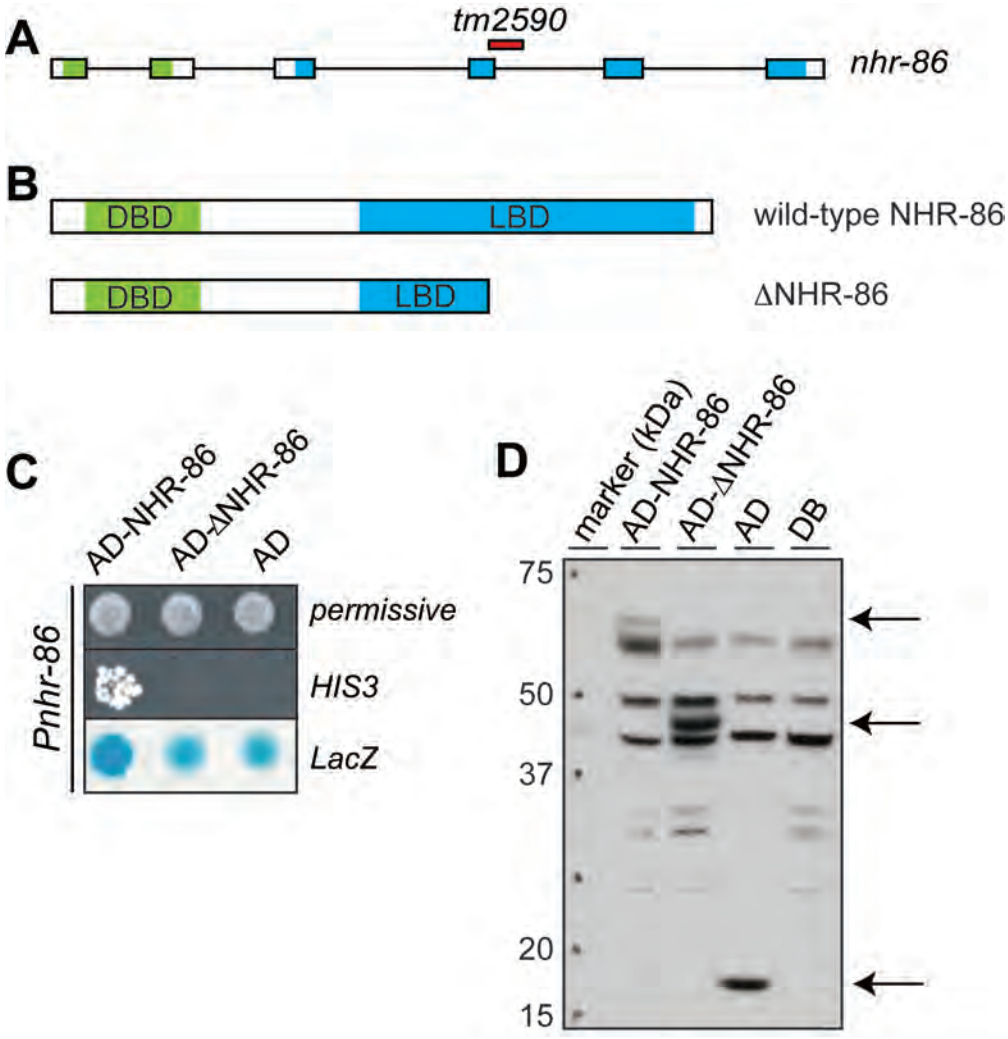

Figure 4.2 *tm2590* is a deletion allele of *nhr-86.*

(A) Gene model of *nhr-86* indicating the *tm2590* deletion. Boxes - exons, lines - introns - red, DNA binding domain - green, and ligand binding domain - blue.

(B) Cartoon illustrating the protein structure of wild-type and the truncated NHR-86, DNA binding domain (DBD), and ligand binding domain (LBD).

(C) Truncated NHR-86 protein fails to bind *Pnhr-86* in Y1H assays: growth on permissive media (top), *HIS3* reporter (middle), and *LacZ* reporter expression (bottom).

(D) Western blot using anti-Gal4AD antibody shows expressed Gal4AD fusions of NHR-86 wild type (AD-NHR-86) and truncated (AD-ΔNHR-86) forms in the *Pnhr-86* Y1H bait (arrows). Extracts from yeast cells harboring Gal4AD (AD) and Gal4DB (DB) vectors with no inserts were used as positive and negative controls respectively.

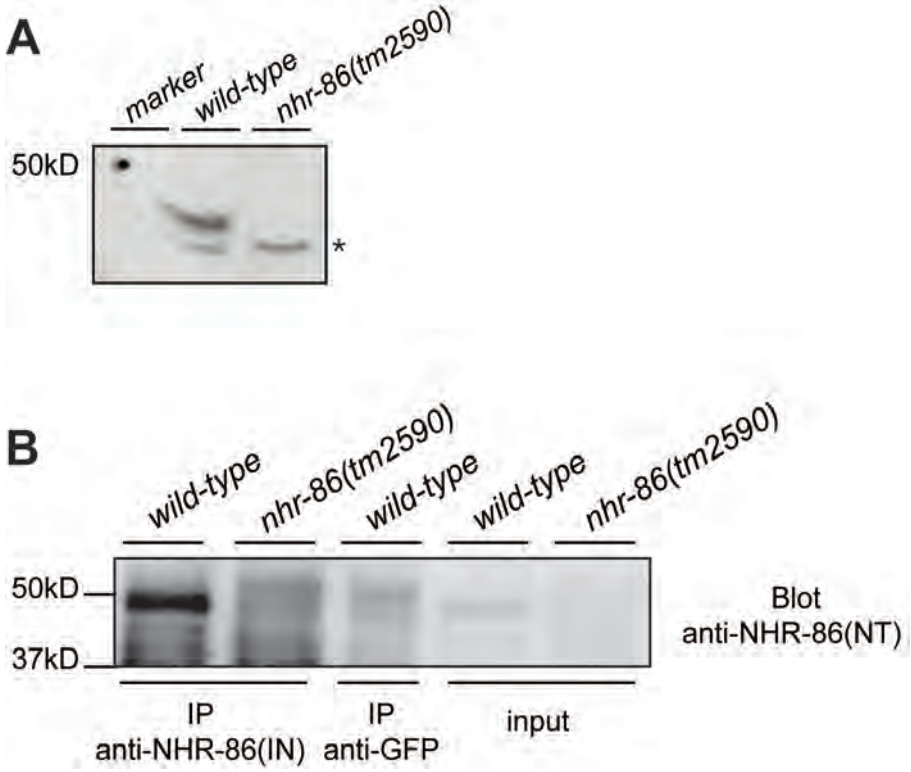

Figure 4.3 *tm2590* animals cannot produce wild type NHR-86 protein

(A) Western blot using anti-NHR-86(NT) antibody shows that *nhr-86(tm2590)* mutant animals do not produce full-length NHR-86 protein (46 kDa): total protein extracts from wild-type, or *nhr-86(tm2590)* mutant animals were used; (asterisk) non-specific band.

(B) Western blot using anti-NHR-86(NT) antibody shows that anti-NHR-86(IN) antibody can specifically immunoprecipitate (IP) NHR-86 protein from a wild-type extract, and the corresponding band in the *nhr-86(tm2590)* extract is absent. Anti-GFP antibody was used as a negative control for the IP experiment.

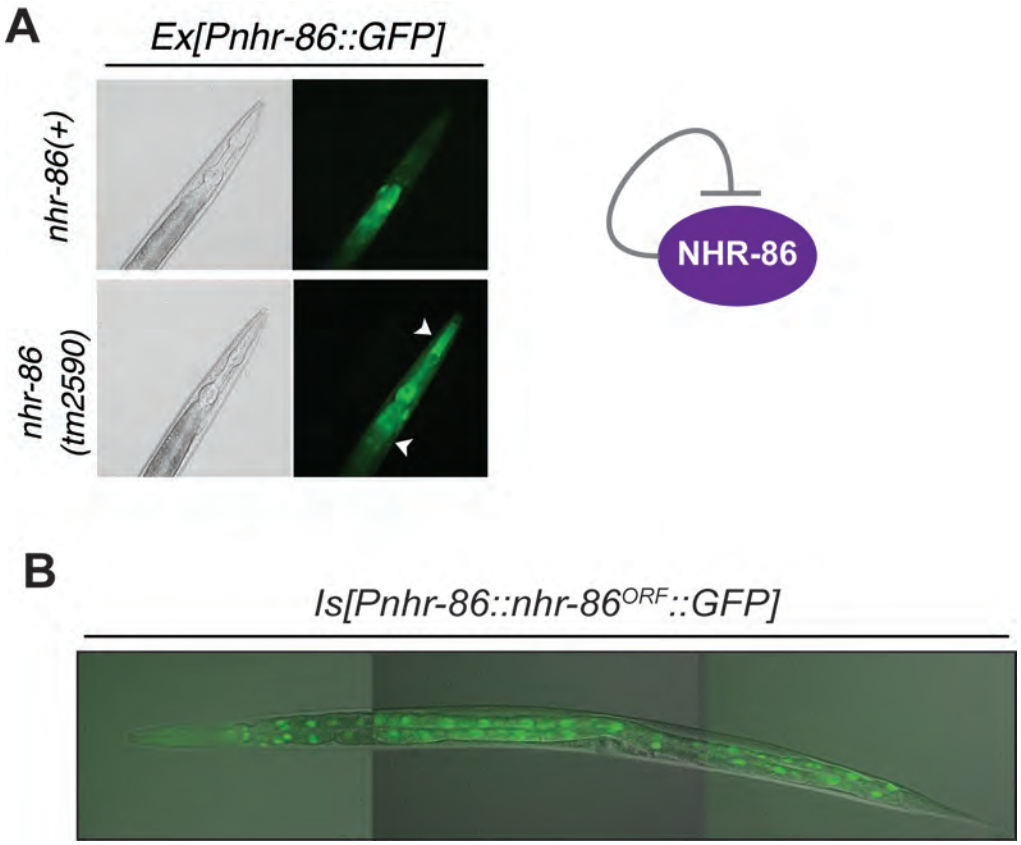

Figure 4.4 NHR-86 is an auto-repressor, and is expressed in metabolic tissues.

(A) *Pnhr-86* activity determined by GFP expression in wild-type (top) or *nhr-86(tm2590)* mutant animals (bottom), indicating auto-repression by NHR-86 in head hypodermis and in the pharynx (white arrowheads).

(B) Expression pattern and subcellular localization of NHR-86 as shown by GFP expression in transgenic animals carrying a *Pnhr-86::nhr-86^{ORF}::GFP* reporter construct. The anterior is to the left.

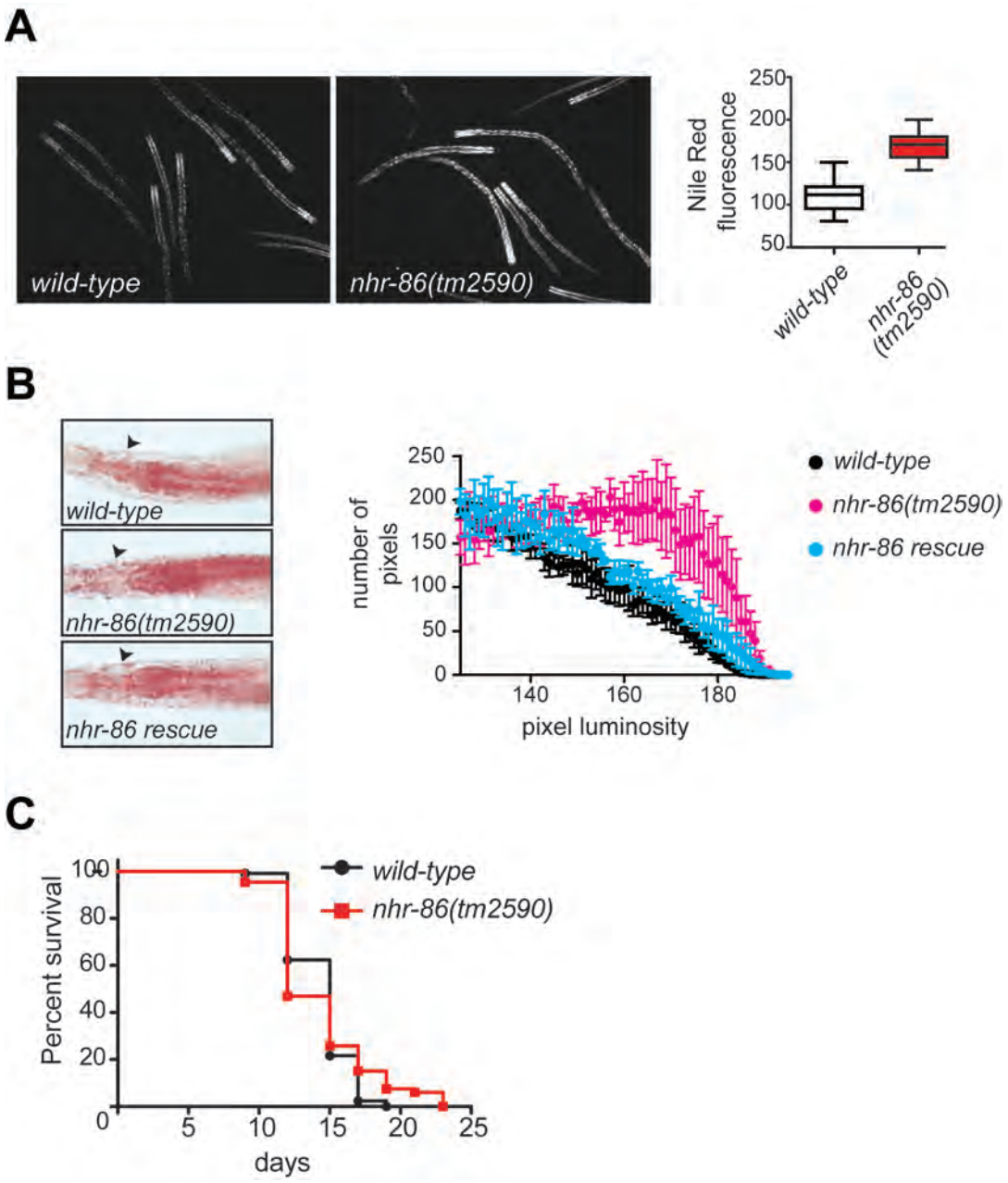

Figure 4.5 *nhr-86(tm2590)* animals have higher amount of body fat.

(A) Nile Red staining of wild-type and *nhr-86(tm2590)* animals (left panel), shows that *nhr-86(tm2590)* animals have increased Nile Red phenotype. The box-plot graph shows the quantification of Nile Red assays as (right) Red box plot indicates changes in Nile Red staining are statistically significant.

(B) Oil-Red-O staining of wild-type, *nhr-86(tm2590)*, and *nhr-86(tm2590)* animals expressing the wildtype transgene (rescue). Black arrowheads point to the posterior bulb of the pharynx. The graph on the right shows the quantification of Oil-Red-O staining assays. *nhr-86(tm2590)* animals accumulate more intensely stained lipid droplets, and this phenotype is rescued by the wild-type *nhr-86* gene. Error bars indicate SEM.

(C) Survival curve of wild-type and *nhr-86(tm2590)* animals shows that loss of *nhr-86* does not affect mean lifespan.

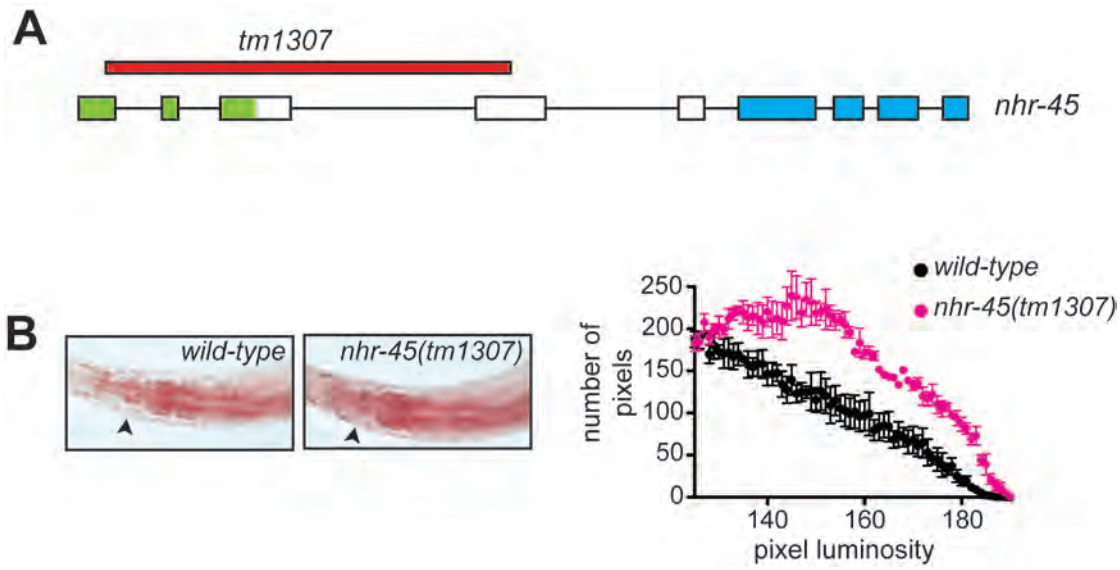

Figure 4.6 *tm1307* is a null allele of *nhr-45.*

(A) Gene model of *nhr-45* depicting the *tm1307* deletion/insertion. Boxes - exons, lines - introns, green - DNA binding domain, blue - ligand binding domain.

(B) Oil-Red-O staining of wild-type and *nhr-45(tm1307)* animals (left). Quantification of the Oil-Red-O staining (right) shows that *nhr-45(tm1307)* animals have more intensely stained lipid droplets. Error bars indicate SEM.

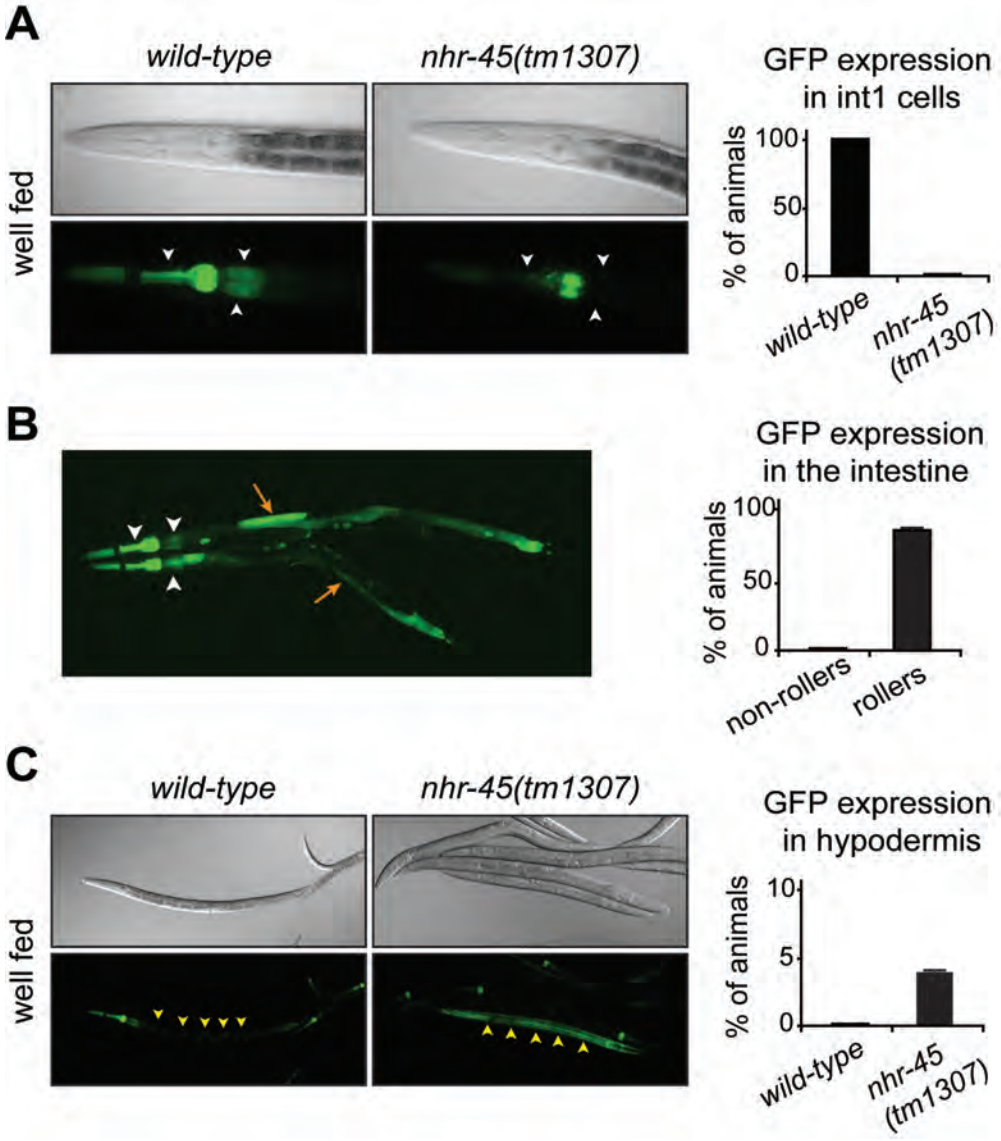

Figure 4.7 *nhr-45* regulates the promoter activity of *nhr-178*.

(A) Images showing *Pnhr-178* activity in wild-type and *nhr-45(tm1307)* animals in the well-fed nutrient state. Graphs next to each group of images show the percentage of animals that exhibit GFP expression in the int1 cells. Error bars indicate the SEM. Top panels - Nomarski images, bottom panels - GFP fluorescence. White arrowheads – first anterior intestinal cells (Int1).

(B) GFP image showing *Pnhr-178* activity when *nhr-45(tm1307)* mutants were injected with a rescuing *nhr-45* genomic fragment along with the dominant co-injection marker *rol-6(su1006),* which causes the animals roll along their long axes (left). Wild-type *nhr-45* restores the GFP expression in the int-1 cells and pharynx (white arrowheads). The rescuing transgene is maintained as a high-copy extrachromosomal array, which results in the mosaic expression in the intestine (orange arrows). Quantification of intestinal GFP expression in the mutant animals that carry the rescuing transgene (rollers) *versus* the ones that don't (non-rollers). Error bars indicate SEM.

(C) Images showing hypodermal *Pnhr-178* activity in *nhr-45(tm1307)* mutants under well-fed conditions, while it is not detected in wild-type animals (yellow arrowheads). Graph on the right shows the percentage of animals that exhibit hypodermal GFP expression. Error bars indicate SEM.

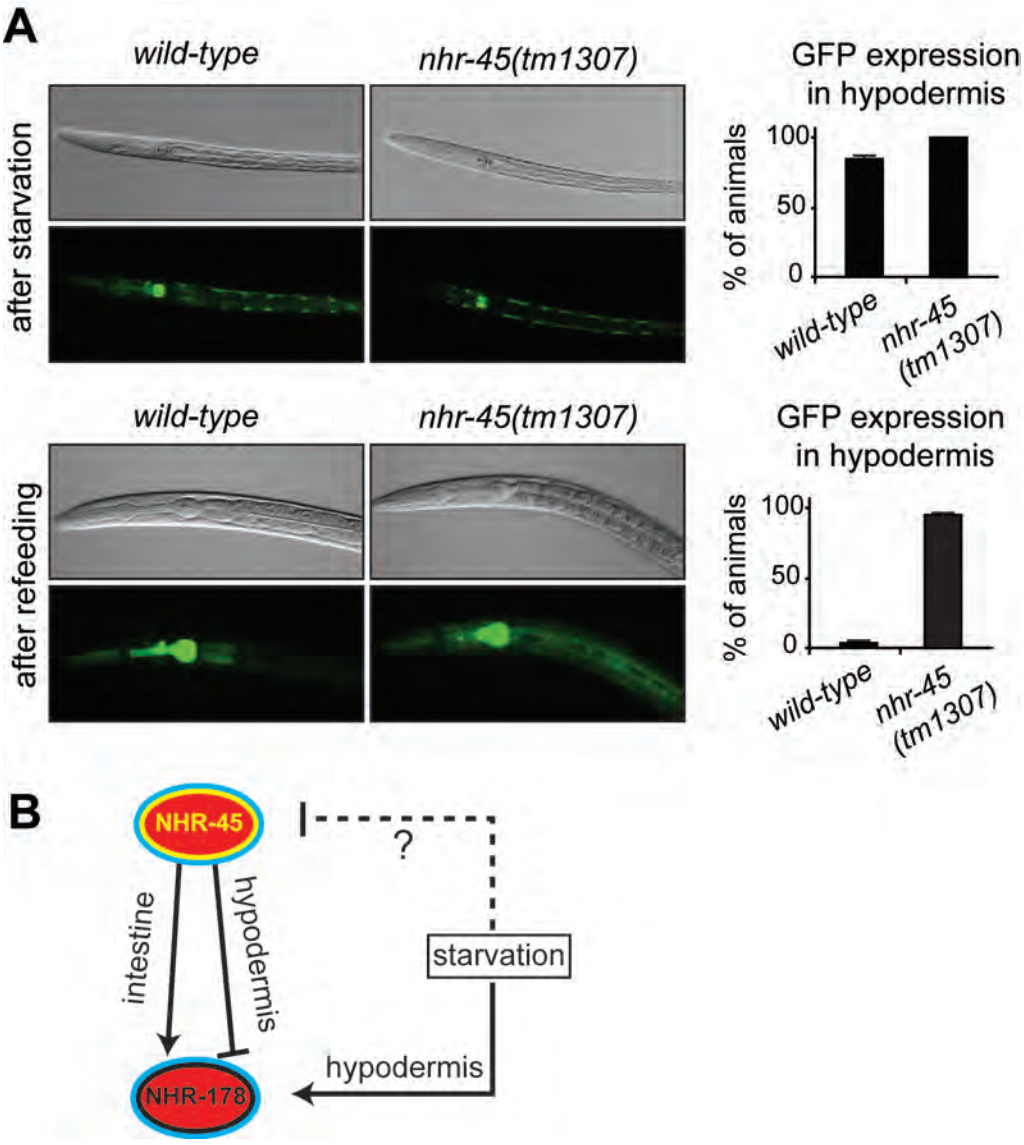

Figure 4.8 An NHR gene circuit that responds to nutrient availability.

(A) Images showing *Pnhr-178* activity in wild-type and *nhr-45(tm1307)* animals in the under different feeding conditions. Graphs next to each group of images show the percentage of animals that exhibit GFP expression in the indicated tissues or cells. Error bars indicate the SEM. Top panels - Nomarski images, bottom panels - GFP fluorescence.

(B) Cartoon depicting effects of NHR-45 and different feeding states on *Pnhr-178* activity.

Preface to Chapter V

Some parts of this chapter have been published separately in:

Arda, H. E., Taubert, S., MacNeil, L. T., Conine, C. C., Tsuda, B., Van Gilst, M., Sequerra, R., Doucette-Stamm, L., Yamamoto, K. R., and Walhout, A. J. M. (2010). Functional modularity of nuclear hormone receptors in a Caenorhabditis elegans metabolic gene regulatory network. Mol. Syst. Biol *6*, 367.

CHAPTER V

Summary and Perspectives

The data presented in the previous chapters provide the first insights into the organizing principles of *C. elegans* GRNs that pertain to systems physiology (Figure 5.1). Our results show that multiple NHRs function in the maintenance of *C. elegans* metabolism, and that they organize into TF modules in the metabolic GRN. We also show that the metabolic cofactor MDT-15 interacts with several of the NHRs that we have identified in the metabolic GRN, and that most of these NHRs confer a lipid staining phenotype.

Below I discuss the broader implications of our findings, and comment on future perspectives about the nematode NHR biology, and physiology.

Expansion of the NHR family in nematodes

After the completion of the *C. elegans* genome sequencing, it was surprising to find that the *C. elegans* genome encodes over 250 predicted NHRs (*C. elegans* Sequencing Consortium, 1998). Phylogenetic analysis of the divergent *C. elegans* NHRs suggests that they share a common HNF4-like ancestor, which had undergone a lineage specific duplication and expansion during the evolution of *C. elegans* (Robinson-Rechavi et al., 2005). In the same study the authors also calculated the evolutionary rates of the divergent *C. elegans* NHRs and found that they evolve much faster than the conserved NHRs. However, even the evolutionary rates of the conserved nematode NHRs are comparably higher than the non-nematode orthologs. Intriguingly, the remarkable

abundance and diversity of NHRs have also been observed in other *Caenorhabditis* species. Haerty *et al.* showed that the genomes of both *Caenorhabditis briggsae* (232 NHRs) and *Caenorhabditis remanei* (256 NHRs) encode similar numbers of NHRs (Haerty et al., 2008). While the authors could identify a three-way reciprocal orthology for about half of the NHRs from the three *Caenorhabditis* species, the remaining NHRs seem to have diverged substantially such that no orthology could be assigned, and are unique to each species. This demonstrates, yet again, the fast evolving feature of the divergent NHRs. Interestingly however, extensive duplication and divergence of the NHR family is not observed in parasitic nematodes [reviewed in (Taubert et al., 2010)]. For instance, the genome of the filarial parasite *Brugia malayi* encodes only 37 NHRs, which is comparable to that of *Drosophila* (21 NHRs) (Ghedin et al., 2007; Scott and Ghedin, 2009). This may suggest that the expansion of the NHR family is a non-parasitic nematode feature, and has been conserved because of the selective pressure on these nematodes in their natural habitats.

C. elegans is a pseudocoelomate with a simple alimentary system composed of the pharynx, intestine and rectum. This is in contrast to animals that have specialized organ systems for digestion, detoxification and endocrine signaling; e.g. liver, kidney and pancreas. *C. elegans* live in soil and need to rapidly respond to changes in nutrient status and environmental conditions, including toxins or pathogens. Along this line, it is interesting to note that in adult mice HNF4 is highly expressed in the liver, kidneys, the adipose tissue and the

gastroenteric system, where HNF4 is necessary for the expression of genes involved in the fatty acid metabolism and detoxification of xenobiotics (Watt et al., 2003; Bookout et al., 2006).

Modularity of NHRs in the C. elegans metabolic GRN

Modularity is an essential component of biological network architecture. Modules consist of nodes that have distinct and separable functions. Yet, each module is not completely independent from the other and they are connected via nodes that have many links or high in-betweenness (Ravasz et al., 2002; Girvan and Newman, 2002). The implications of modularity in biological networks are that it increases the robustness and adaptability of the system (Babu et al., 2004). The discovery of modularity in biological networks spurred researchers to find out how modularity could have emerged at the first place [reviewed in (Wagner et al., 2007)]. One theory suggests that modularity arises by gene duplications followed by mutations that allow new genes to gain or lose links, which under certain parameters eventually leads to modularity. However, this theory does not take natural selection into account and therefore was later claimed inadequate because it fails to explain how modularity is maintained once acquired (Wagner et al., 2007). As an alternative, by using simulation algorithms, Alon and colleagues showed that spontaneous evolution of modularity can occur in response to ‘modularly varying environment’ (Kashtan and Alon, 2005). In other words, when a system is presented with varying ‘subproblems’ of a

common goal, rather than randomly changing goals, networks with modular topology were better at finding a solution. By simply rewiring existing modules that can carry out specific tasks, the system can adapt faster to changing environments without making major alterations in the network structure. In addition, when they analyzed reconstructed metabolic networks of different bacterial species, they found that there is a correlation between the network topology and their habitats; the more variable the environment, the more modular the networks are (Parter et al., 2007).

Our results show that the *C. elegans* metabolic GRN is highly modular, and at least two of these modules consist mainly of NHRs. It is interesting to think that similar environmental pressures might have pushed *C. elegans* metabolic GRNs to acquire this modular structure. Thus, by expanding the NHR family, which is known to act as metabolic sensors and organizing them into TF modules, *C. elegans* metabolic GRNs have a dual mechanism to rapidly respond to physiological or environmental cues.

It is hard to conclude with absolute confidence that nematodes are constantly subjected to changing conditions in their natural habitats, but we can speculate that our findings are consistent with the idea that higher modularity allows higher adaptability and increased fitness. As we continue mapping the GRNs pertaining to *C. elegans* systems physiology, we will likely discover additional modules involving other NHRs, gain more insights into the global architecture of these networks. Together with the functional characterization of

these NHRs we can start addressing how the existence of TF modules and their wiring in the network increase the adaptability of the nematodes to the changing environmental conditions.

Biological functions of the divergent C. elegans *NHRs*

The finding that the *C. elegans* genome encodes for unprecedented numbers of NHRs initially led to the speculations that majority of these divergent NHRs might not be expressed or could be pseudogenes. However, in the years following the genome sequencing, several studies showed that, to the contrary, most *C. elegans* NHRs are expressed and believed to be functional (Miyabayashi et al., 1999; Sluder et al., 1999; Lamesch et al., 2004; Baugh et al., 2009). To date, detailed phenotypic characterizations have been mainly limited to the conserved NHRs, and the functional analysis of divergent NHRs has been challenging mostly due to the overwhelming numbers.

Our study demonstrates that several divergent *C. elegans* NHRs are essential for normal lipid metabolism and nutrient response, and that they are not necessarily functionally redundant. Supporting this notion, in a recent study Baugh *et al*. compared the gene expression profiles of L1 larvae under fed and fasted conditions using high-density microarrays and found that the transcript levels of numerous NHRs changed significantly, suggesting that the nematode metabolic system employs multiple NHRs to maintain its energy balance and survival (Baugh et al., 2009).

In the future, it will be important to obtain more information about the biological roles of the divergent *C. elegans* NHRs by designing systems-level assays to systematically address their functionality. For example, analysis of the spatio-temporal expression pattern of the remaining *C. elegans* NHRs may be informative to understand the evolution of gene function. Do these NHRs all have distinct or mostly similar expression patterns, and how does that relate to their position in the phylogenetic tree? Other functional assays may include perturbing the 'normal environment' with xenobiotics, metabolic and oxidative stress, or pathogens and analyze the contribution of NHRs to the fitness of the nematode under these conditions.

Extent of dimerization of C. elegans *NHRs*

NHRs can bind to their DNA targets in the form of homo or heterodimers. Heterodimerization is a molecular strategy to increase DNA binding site complexity and the diversity of functional TFs without increasing the number of individual components (Grove and Walhout, 2008; Grove et al., 2009). For instance, the mammalian nuclear receptor RXR functions as a heterodimerization partner/hub for large number of other mammalian NHRs (Mangelsdorf and Evans, 1995). Being a liganded nuclear receptor itself, RXR heterodimerization adds another layer of complexity and versatility to hormone signaling pathways, although ligand dependent RXR activation in these heterodimer complexes is context dependent. When RXR is dimerized to retinoic acid or vitamin D nuclear

receptors, RXR ligands alone cannot induce transcription and RXR is considered to be a ‘silent’ partner. In contrast, when partnered with LXR, FXR or PPARs, ligand induced activation of RXR is permitted, and these NHRs become transcriptionally active, even though they are not bound by their cognate ligands (Blumberg and Evans, 1998; Gronemeyer et al., 2004).

As mentioned earlier in Chapter I, despite having 284 NHRs, the *C. elegans* genome does not have an RXR ortholog, even though RXRs belong to the evolutionarily ancient NR2B class (Escriva et al., 2000). In addition, HNF4 nuclear receptors, from which most *C. elegans* NHRs have expanded, are only known to homodimerize (Sladek et al., 1990; Watt et al., 2003). This raises the question of whether in *C. elegans* NHRs work predominantly as monomers or homodimers. Given the high evolutionary rates of divergent *C. elegans* NHRs, it is also thought that the requirement of an obligate dimerization partner to become transcriptionally active is a less plausible evolutionary scenario than functioning as a homodimer (Robinson-Rechavi et al., 2005). Nevertheless, studies using high-throughput Y2H assays suggest that at least a few *C. elegans* NHRs can interact with several other NHRs, and may function as heterodimer hubs (Li et al., 2004; Vermeirssen et al., 2007). For example, NHR-49, an HNF4-like NHR, dimerizes with 25 NHRs, and in Chapter II we showed that at least three of them have a fat storage phenotype when knocked down by RNAi. It is possible that the extensive divergence of HNF4-like *C. elegans* NHRs could have evolved novel and unique contact interfaces for heterodimerization that do not

exist in the ancestor. Systematic analysis of NHR-NHR interactions would be essential to determine the extent of NHR dimerization in the nematode, and to further our understanding of the role of NHRs in *C. elegans* physiology.

Ligand binding ability of C. elegans *NHRs*

Perhaps the most important feature of NHRs is their ability to interact with diffusible ligands, which can modulate their transcriptional output. Due to the therapeutic potential, substantial effort had been made to identify ligands of the mammalian nuclear receptors and understand the molecular basis of ligand binding. Currently, ligands for more than half of the human NHRs have been identified, and several synthetic ligands are being tested in clinical trials for the treatment of diabetes, obesity and hypertension (Gronemeyer et al., 2004; Sonoda et al., 2008).

Despite the abundance of *C. elegans* NHRs, thus far, only the ligand for the vitamin D receptor ortholog DAF-12 has been identified (Motola et al., 2006). DAF-12 regulates several aspects of nematode physiology, including fat metabolism, developmental timing and adult lifespan (Magner and Antebi, 2008). Under favorable conditions, the *C. elegans* neuroendocrine pathways trigger the production of *dafachronic acid*, the DAF-12 ligand, to promote reproductive larval fate. Conversely, when the environment is unfavorable, the dafachronic acid production is believed to be repressed, and in the absence of its ligand, DAF-12

associates with its corepressor DIN-1, which pushes the nematodes into the dauer state (Magner and Antebi, 2008).

Dafachronic acid is a cholesterol derivative, and an endogenously synthesized hormone (Motola et al., 2006). The discovery of these hormone synthesis pathways in *C. elegans* revealed the fact that they are conserved from nematodes to humans. Combined with powerful genetics, studies of the *C. elegans* hormone pathways may provide insights into the regulation of steroid signaling.

DAF-12 does not belong to the group of divergent HNF4-like NHRs, and it remains to be determined whether the rest of the *C. elegans* NHRs are liganded or not. Initial phylogenetic analyses argued that ancient NHRs are mostly orthologs of orphan receptors and in case of the mammalian NHRs, there is no relationship between the position of the NHRs in the phylogenetic tree and the chemical nature of the ligands that they bind (Escriva et al., 2000; Gronemeyer et al., 2004). Therefore, it was suggested that the first nuclear receptor was likely not liganded and that ligand binding was acquired independently several times during the course of NHR evolution. Certainly, phylogenetic analyses are limited by the availability genome sequences, and the thoughts on the origins of liganded nuclear receptors are anticipated to change as researchers continue to complete sequencing the genomes of diverse species. In addition, the fact that no ligand has been identified for the ‘orphan receptors’ should not necessarily lead to the conclusion that they are truly incapable of binding any small molecule.

For instance, the mammalian HNF4 was believed to be an orphan receptor until it was co-crystallized with endogenous fatty acids inside its ligand binding domain (Wisely et al., 2002).

Considering the biological roles of the *C. elegans* NHRs, which include molting, fat metabolism, longevity, reproduction and nutrient sensing, it is reasonable to speculate that *C. elegans* NHRs other than DAF-12 can also bind specific ligands. Interestingly, there is less sequence conservation in the ligand binding domain than in the DNA binding domain among the *C. elegans* NHRs (Sluder and Maina, 2001; Robinson-Rechavi et al., 2005). This indicates that perhaps expression of numerous NHRs with divergent ligand binding domains gives the nematodes the ability to recognize and respond to a variety of signals and provides a better adaptability to their environment. Future experiments will uncover additional ligands and reveal whether each NHR responds to a single ligand, or if there may be redundancy or even modularity at this level as well.

DNA binding specificity of C. elegans *NHRs*

NHRs bind to hormone response elements through the 'P-box' region, a conserved string of amino acid sequence that exists in the DNA binding domain (Zilliacus et al., 1995). The P-box sequences of several divergent *C. elegans* NHRs are very different than those of the vertebrates, and most of them are only found in nematodes (Van Gilst et al., 2002). Therefore, predicting the DNA binding specificity of the *C. elegans* NHRs based on P-box sequence

conservation alone is not going to be feasible. Thus far, no ChIP data is available for any of the *C. elegans* NHRs and DAF-12 remains as the only nematode NHR whose DNA binding motif has been experimentally characterized (Shostak et al., 2004).

Y1H assays provide a valuable resource for the identification of TF binding sites. Previously, by using our Y1H data and a combination of available computational algorithms, we have successfully delineated the consensus binding site for several *C. elegans* TFs from different families, and showed that these binding sites are both necessary and sufficient for the binding event to occur (Deplancke et al., 2006; Ow et al., 2008; Reece-Hoyes et al., 2009). In this study, by using Y1H assays, we have identified promoter targets for 27 *C. elegans* NHRs, which is approximately 10% of all the NHRs in the nematode. This demonstrates the power of gene-centered network mapping. Further analyses of these promoter fragments will be instrumental in determining the DNA binding specificities of the *C. elegans* NHRs that occur in the metabolic GRN.

Cofactor-NHR interactions

Transcriptional cofactors are important regulators of mammalian physiology, and several cofactor complexes have been discovered that work in conjunction with NHRs (Rosenfeld et al., 2006). Biochemical analyses showed that one of the key structural determinants of the cofactor-NHR interactions is the

presence of a conserved LXXLL motif in the cofactors. Often, this motif enables vertebrate cofactors to interact with the ligand binding domain of the NHRs, either in the presence or absence of their cognate ligands depending on the cellular context.

Initial survey of the *C. elegans* genome could not identify orthologous cofactors proteins with known LXXLL cofactor-NHR interfaces (Taubert et al., 2006; Näär and Thakur, 2009). In fact, most of the well-characterized vertebrate cofactors, such as PGC-1α, and the p160 family of coactivators are absent from the yeast, worm and fruitfly genomes. Nevertheless, MDT-15 and DIN-1 are two examples of *C. elegans* transcriptional cofactors that can interact with NHRs to regulate important physiological processes, suggesting that cofactor-NHR interactions are ancient mechanisms for metabolic gene regulation. Given the sequence divergence of the ligand binding domains of the nematode NHRs, it is possible that they might employ different regulatory mechanisms and novel interaction surfaces to interact with cofactors. It is also likely that cofactors that are unique to the nematodes might exist, complementing the expansion of the NHR lineage. The use of systems-level approaches, such as high-throughput Y2H assays, to systematically identify cofactor-NHR interactions will be pivotal for reconstructing and expanding on current metabolic gene regulatory networks, and undoubtedly will advance our understanding of the evolution of cofactor:NHR interactions.

Final word

In this chapter, I sought to discuss the broader implications of our findings, and comment on future perspectives about the nematode NHR biology, and nematode physiology. The networks described in this study are only the first steps to start experimentally delineating metazoan metabolic GRNs. It is necessary to continue our efforts to get a better understanding of the *C. elegans* NHR biology. Identification of ligands, dimerization networks, and cofactor interactions will ultimately help us comprehend the complexities of the physiological networks. Despite the considerable progress that had been made using vertebrate models, studies of *C. elegans* systems physiology provide a unique opportunity to investigate the evolution of metazoan metabolic GRNs, and discover the novel mechanisms that organisms invent to address common problems, such as energy balance and homeostasis.

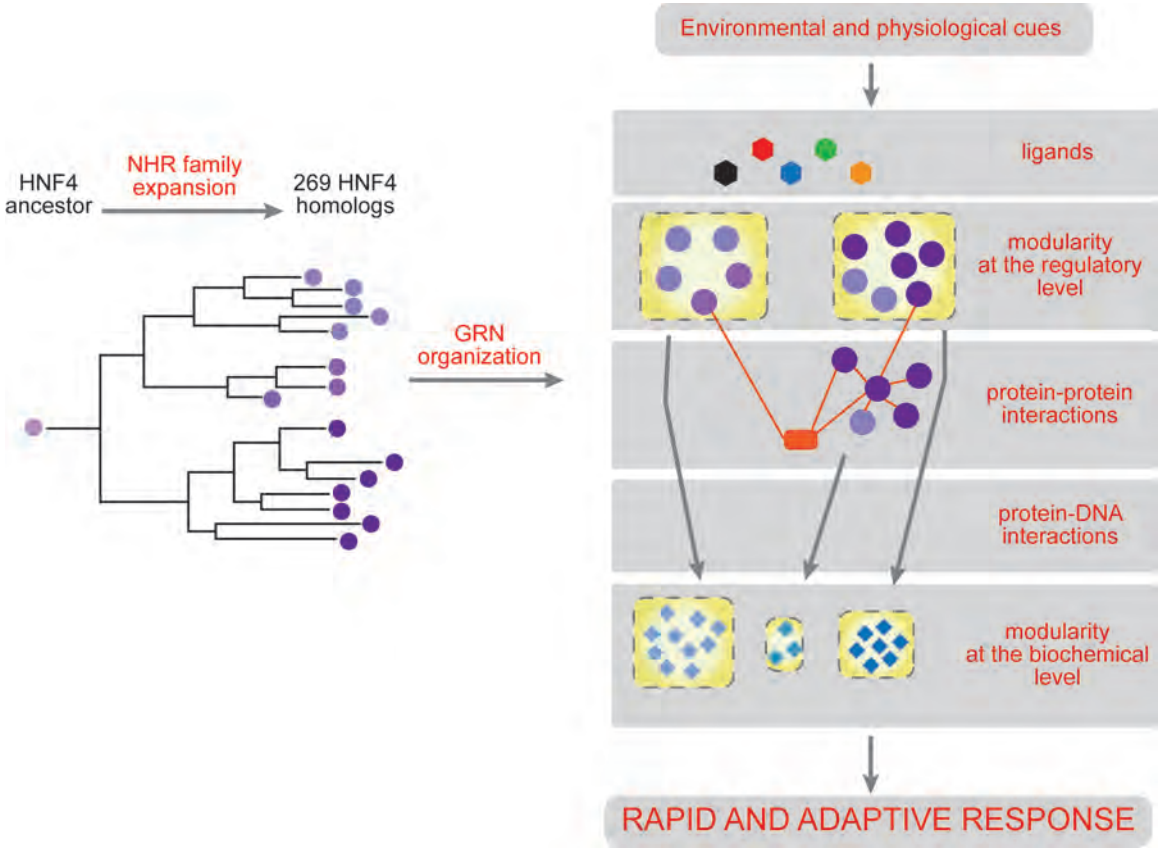

Figure 5.1 Model for the organizing principles of *C. elegans* metabolic networks.

Circles are NHRs, and hues of purple represent the evolution of HNF4 family NHRs in *C. elegans*. These NHRs are organized into TF modules in the metabolic GRN. NHRs regulate their metabolic target genes (blue diamonds) by interacting with ligands (hexagons), and with other proteins (orange lines) such as the cofactor MDT-15 (orange oval) and dimerization partners.

APPENDIX

List of *C. elegans* strains generated in this study

Strain Name	*Genotype*[*]
VL220	*unc-119(ed3) III; wwEx52[Pnhr-86::GFP unc-119(+)]*
VL484	*nhr-45(tm1307) X, (x4)*
VL491	*nhr-86(tm2590) V, (x6)*
VL505	*unc-119(ed3) III; wwIs22[Pnhr-86::nhr-86^{ORF}::GFP unc-119(+)]*
VL510	*nhr-86(tm2590) V; wwEx52[Pnhr-86::GFP unc-119(+)]*
VL584[†]	*nhr-86(tm2590) V; wwIs22[Pnhr-86::nhr-86ORF::GFP unc-119(+)]*
VL600[‡]	*unc-119(ed3) III; wwIs23[Pnhr-178::GFP unc-119(+)]*
VL714	*unc-119(ed3) III; wwEx53[Pacdh-2::GFP]*
VL717	*unc-119(ed3) III; wwEx54[Pacdh-1::GFP]*
VL739	*nhr-45(tm1307) X; wwIs23[Pnhr-178::GFP unc-119(+)]*
VL752	*nhr-45(tm1307) X; wwIs23[Pnhr-178::GFP unc-119(+)]; wwEx55[Pnhr-45::nhr-45^{genomic}::nhr-45 3'UTR rol-6(su1006)]*

[*] all strains were generated using out-crossed alleles, and the number of out-crosses is indicated after each allele name

[†] transgene is integrated into the genome of the nematodes

[‡] transgene is integrated into the genome of the nematodes

BIBLIOGRAPHY

Alon, U. (2007). Network motifs: theory and experimental approaches. Nat Rev Genet *8*, 450-461.

Ambros, V. (2004). The functions of animal microRNAs. Nature *431*, 350-355.

Arda, H. E., Taubert, S., MacNeil, L. T., Conine, C. C., Tsuda, B., Van Gilst, M., Sequerra, R., Doucette-Stamm, L., Yamamoto, K. R., and Walhout, A. J. M. (2010). Functional modularity of nuclear hormone receptors in a *Caenorhabditis elegans* metabolic gene regulatory network. Mol Syst Biol *6,* 367.

Arda, H. E., and Walhout, A. J. (2009). Gene-centered regulatory networks. Brief Funct Genomic Proteomic *9*, 4-12.

Ashrafi, K. (2007). Obesity and the regulation of fat metabolism. WormBook, 1-20.

Ashrafi, K., Chang, F. Y., Watts, J. L., Fraser, A. G., Kamath, R. S., Ahringer, J., and Ruvkun, G. (2003). Genome-wide RNAi analysis of Caenorhabditis *elegans* fat regulatory genes. Nature *421*, 268-272.

Babu, M. M., Luscombe, N. M., Aravind, L., Gerstein, M., and Teichmann, S. A. (2004). Structure and evolution of transcriptional regulatory networks. Curr. Opin. Struct. Biol *14*, 283-291.

Barabasi, and Albert (1999). Emergence of scaling in random networks. Science *286*, 509-512.

Barabási, A., and Oltvai, Z. N. (2004). Network biology: understanding the cell's functional organization. Nat. Rev. Genet *5*, 101-113.

Barrasa, M. I., Vaglio, P., Cavasino, F., Jacotot, L., and Walhout, A. J. M. (2007). EDGEdb: a transcription factor-DNA interaction database for the analysis of *C. elegans* differential gene expression. BMC Genomics *8*, 21.

Barrett, C. L., Herring, C. D., Reed, J. L., and Palsson, B. O. (2005). The global transcriptional regulatory network for metabolism in Escherichia coli exhibits few dominant functional states. Proceedings of the

National Academy of Sciences of the United States of America *102*, 19103-19108.

Baugh, L. R., DeModena, J., and Sternberg, P. W. (2009). RNA Pol II Accumulates at Promoters of Growth Genes During Developmental Arrest. Science *324*, 92-94.

Blumberg, B., and Evans, R. M. (1998). Orphan nuclear receptors—new ligands and new possibilities. Genes & Development *12*, 3149-3155.

Bookout, A. L., Jeong, Y., Downes, M., Yu, R. T., Evans, R. M., and Mangelsdorf, D. J. (2006). Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network. Cell *126*, 789-799.

Borneman, A. R., Leigh-Bell, J. A., Yu, H., Bertone, P., Gerstein, M., and Snyder, M. (2006). Target hub proteins serve as master regulators of development in yeast. Genes & Development *20*, 435-448.

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

Brown, M. S., and Goldstein, J. L. (1997). The SREBP Pathway: Regulation of Cholesterol Metabolism by Proteolysis of a Membrane-Bound Transcription Factor. Cell *89*, 331-340.

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

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green fluorescent protein as a marker for gene expression. Science *263*, 802-805.

Chawla, A., Repa, J. J., Evans, R. M., and Mangelsdorf, D. J. (2001). Nuclear receptors and lipid physiology: opening the X-files. Science *294*, 1866-1870.

Davidson, E. H., Rast, J. P., Oliveri, P., Ransick, A., Calestani, C., Yuh, C., Minokawa, T., Amore, G., Hinman, V., Arenas-Mena, C., et al. (2002). A Genomic Regulatory Network for Development. Science *295*, 1669-1678.

Deplancke, B., Dupuy, D., Vidal, M., and Walhout, A. J. M. (2004). A gateway-compatible yeast one-hybrid system. Genome Res *14*, 2093-2101.

Deplancke, B., Mukhopadhyay, A., Ao, W., Elewa, A. M., Grove, C. A., Martinez, N. J., Sequerra, R., Doucette-Stamm, L., Reece-Hoyes, J. S., Hope, I. A., et al. (2006a). A gene-centered *C. elegans* protein-DNA interaction network. Cell *125*, 1193-1205.

Deplancke, B., Vermeirssen, V., Arda, H. E., Martinez, N. J., and Walhout, A. J. M. (2006b). Gateway-compatible yeast one-hybrid screens.

Dupuy, D., Li, Q., Deplancke, B., Boxem, M., Hao, T., Lamesch, P., Sequerra, R., Bosak, S., Doucette-Stamm, L., Hope, I. A., et al. (2004). A first version of the Caenorhabditis *elegans* Promoterome. Genome Res *14*, 2169-2175.

Erwin, D. H., and Davidson, E. H. (2009). The evolution of hierarchical gene regulatory networks. Nat Rev Genet *10*, 141-148.

Escriva, H., Delaunay, F., and Laudet, V. (2000). Ligand binding and nuclear receptor evolution. BioEssays *22*, 717-727.

Ge, K., Guermah, M., Yuan, C., Ito, M., Wallberg, A. E., Spiegelman, B. M., and Roeder, R. G. (2002). Transcription coactivator TRAP220 is required for PPAR[gamma]2-stimulated adipogenesis. Nature *417*, 563-567.

Gehring, W. J., and Ikeo, K. (1999). Pax 6: mastering eye morphogenesis and eye evolution. Trends Genet *15*, 371-377.

Ghedin, E., Wang, S., Spiro, D., Caler, E., Zhao, Q., Crabtree, J., Allen, J. E., Delcher, A. L., Guiliano, D. B., Miranda-Saavedra, D., et al. (2007). Draft genome of the filarial nematode parasite Brugia malayi. Science *317*, 1756-1760.

Girvan, M., and Newman, M. E. J. (2002). Community structure in social and biological networks. Proceedings of the National Academy of Sciences of the United States of America *99*, 7821-7826.

Gronemeyer, H., Gustafsson, J., and Laudet, V. (2004). Principles for modulation of the nuclear receptor superfamily. Nat Rev Drug Discov *3*, 950-964.

Grove, C. A., De Masi, F., Barrasa, M. I., Newburger, D. E., Alkema, M. J., Bulyk, M. L., and Walhout, A. J. M. (2009). A multiparameter network reveals extensive divergence between *C. elegans* bHLH transcription factors. Cell *138*, 314-327.

Grove, C. A., and Walhout, A. J. M. (2008). Transcription factor functionality and transcription regulatory networks. Mol Biosyst *4*, 309-314.

Gunsalus, K. C., Ge, H., Schetter, A. J., Goldberg, D. S., Han, J. J., Hao, T., Berriz, G. F., Bertin, N., Huang, J., Chuang, L., et al. (2005). Predictive models of molecular machines involved in Caenorhabditis *elegans* early embryogenesis. Nature *436*, 861-865.

Haerty, W., Artieri, C., Khezri, N., Singh, R. S., and Gupta, B. P. (2008). Comparative analysis of function and interaction of transcription factors in nematodes: extensive conservation of orthology coupled to rapid sequence evolution. BMC Genomics *9*, 399.

Harbison, C. T., Gordon, D. B., Lee, T. I., Rinaldi, N. J., Macisaac, K. D., Danford, T. W., Hannett, N. M., Tagne, J., Reynolds, D. B., Yoo, J., et al. (2004). Transcriptional regulatory code of a eukaryotic genome. Nature *431*, 99-104.

Hayhurst, G. P., Lee, Y. H., Lambert, G., Ward, J. M., and Gonzalez, F. J. (2001). Hepatocyte nuclear factor 4alpha (nuclear receptor 2A1) is essential for maintenance of hepatic gene expression and lipid homeostasis. Mol. Cell. Biol *21*, 1393-1403.

Huang, P., Chandra, V., and Rastinejad, F. (2010). Structural Overview of the Nuclear Receptor Superfamily: Insights into Physiology and Therapeutics. Annu. Rev. Physiol. *72*, 247-272.

Hunt-Newbury, R., Viveiros, R., Johnsen, R., Mah, A., Anastas, D., Fang, L., Halfnight, E., Lee, D., Lin, J., Lorch, A., et al. (2007). High-Throughput In Vivo Analysis of Gene Expression in Caenorhabditis *elegans*. PLoS Biol *5*, e237.

Ihmels, J., Levy, R., and Barkai, N. (2004). Principles of transcriptional control in the metabolic network of *Saccharomyces cerevisiae*. Nat Biotech *22*, 86-92.

International Human Genome Sequencing Consortium (2004). Finishing the euchromatic sequence of the human genome. Nature *431*, 931-945.

Kamath, R. S., Fraser, A. G., Dong, Y., Poulin, G., Durbin, R., Gotta, M., Kanapin, A., Le Bot, N., Moreno, S., Sohrmann, M., et al. (2003). Systematic functional analysis of the *Caenorhabditis elegans* genome using RNAi. Nature *421*, 231-237.

Kashtan, N., and Alon, U. (2005). Spontaneous evolution of modularity and network motifs. Proceedings of the National Academy of Sciences of the United States of America *102*, 13773-13778.

Kim, T. H., and Ren, B. (2006). Genome-wide analysis of protein-DNA interactions. Annu Rev Genomics Hum Genet *7*, 81-102.

Lamesch, P., Milstein, S., Hao, T., Rosenberg, J., Li, N., Sequerra, R., Bosak, S., Doucette-Stamm, L., Vandenhaute, J., Hill, D. E., et al. (2004). *C. elegans* ORFeome Version 3.1: Increasing the Coverage of ORFeome Resources With Improved Gene Predictions. Genome Research *14*, 2064-2069.

Lehner, B., and Lee, I. (2008). Network-guided genetic screening: building, testing and using gene networks to predict gene function. Briefings in Functional Genomics and Proteomics *7*, 217-227.

Levine, M., and Davidson, E. H. (2005). Gene regulatory networks for development. Proceedings of the National Academy of Sciences of the United States of America *102*, 4936-4942.

Li, J. J., and Herskowitz, I. (1993). Isolation of ORC6, a component of the yeast origin recognition complex by a one-hybrid system. Science *262*, 1870-1874.

Li, S., Armstrong, C. M., Bertin, N., Ge, H., Milstein, S., Boxem, M., Vidalain, P., Han, J. J., Chesneau, A., Hao, T., et al. (2004). A map of the interactome network of the metazoan *C. elegans*. Science *303*, 540-543.

Li, X., MacArthur, S., Bourgon, R., Nix, D., Pollard, D. A., Iyer, V. N., Hechmer, A., Simirenko, L., Stapleton, M., Luengo Hendriks, C. L., et al. (2008). Transcription factors bind thousands of active and inactive regions in the Drosophila blastoderm. PLoS Biol *6*, e27.

Lin, J. D., Lin, I., H, C., schin, and Spiegelman, B. M. (2005). Metabolic control through the PGC-1 family of transcription coactivators. Cell Metabolism *1*, 361-370.

Ludewig, A. H., Kober-Eisermann, C., Weitzel, C., Bethke, A., Neubert, K., Gerisch, B., Hutter, H., and Antebi, A. (2004). A novel nuclear receptor/coregulator complex controls *C. elegans* lipid metabolism, larval development, and aging. Genes Dev *18*, 2120-2133.

Maduro, M. F., and Rothman, J. H. (2002). Making worm guts: the gene regulatory network of the *Caenorhabditis elegans* endoderm. Dev. Biol *246*, 68-85.

Maglich, J., Sluder, A., Guan, X., Shi, Y., McKee, D., Carrick, K., Kamdar, K., Willson, T., and Moore, J. (2001). Comparison of complete nuclear receptor sets from the human, Caenorhabditis *elegans* and Drosophila genomes. Genome Biology *2*, research0029.1-research0029.7.

Magner, D. B., and Antebi, A. (2008). Caenorhabditis *elegans* nuclear receptors: insights into life traits. Trends in Endocrinology & Metabolism *19*, 153-160.

Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. Proc. Natl. Acad. Sci. U.S.A *100*, 11980-11985.

Mangelsdorf, D. J., and Evans, R. M. (1995). The RXR heterodimers and orphan receptors. Cell *83*, 841-850.

Maniatis, T., Fritsch, E., Sambrook, J., and Russell, D. (1982). Molecular Cloning: A Laboratory Manual Second edition. (Cold Spring Harbor Laboratory Press).

Martinez, N. J., Ow, M. C., Barrasa, M. I., Hammell, M., Sequerra, R., Doucette-Stamm, L., Roth, F. P., Ambros, V. R., and Walhout, A. J. M. (2008). A *C. elegans* genome-scale microRNA network contains composite feedback motifs with high flux capacity. Genes Dev *22*, 2535-2549.

Martinez, N. J., and Walhout, A. J. M. (2009). The interplay between transcription factors and microRNAs in genome-scale regulatory networks. Bioessays *31*, 435-445.

Martinez, N. J., Ow, M. C., Reece-Hoyes, J. S., Barrasa, M. I., Ambros, V. R., and Walhout, A. J. (2008). Genome-scale spatiotemporal analysis of Caenorhabditis *elegans* microRNA promoter activity. Genome Research *18*, 2005-2015.

McGhee, J. (2007). The *C. elegans* intestine. WormBook. Available at: http://www.wormbook.org/chapters/www_intestine/intestine.html [Accessed June 9, 2010].

McKay, R. M., McKay, J. P., Avery, L., and Graff, J. M. (2003). C *elegans*: a model for exploring the genetics of fat storage. Dev. Cell *4*, 131-142.

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

Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D., and Alon, U. (2002). Network motifs: simple building blocks of complex networks. Science *298*, 824-827.

Miyabayashi, T., Palfreyman, M. T., Sluder, A. E., Slack, F., and Sengupta, P. (1999). Expression and Function of Members of a Divergent Nuclear Receptor Family in Caenorhabditis *elegans*. Developmental Biology *215*, 314-331.

Motola, D. L., Cummins, C. L., Rottiers, V., Sharma, K. K., Li, T., Li, Y., Suino-Powell, K., Xu, H. E., Auchus, R. J., Antebi, A., et al. (2006). Identification of Ligands for DAF-12 that Govern Dauer Formation and Reproduction in *C. elegans*. Cell *124*, 1209-1223.

Mukherjee, S., Berger, M. F., Jona, G., Wang, X. S., Muzzey, D., Snyder, M., Young, R. A., and Bulyk, M. L. (2004). Rapid analysis of the DNA-binding specificities of transcription factors with DNA microarrays. Nat. Genet *36*, 1331-1339.

Mullaney, B. C., and Ashrafi, K. (2009). *C. elegans* fat storage and metabolic regulation. Biochimica et Biophysica Acta (BBA) - Molecular and Cell Biology of Lipids *1791*, 474-478.

Näär, A. M., Lemon, B. D., and Tjian, R. (2001). Transcriptional coactivator complexes. Annu. Rev. Biochem *70*, 475-501.

Näär, A. M., and Thakur, J. K. (2009). Nuclear receptor-like transcription factors in fungi. Genes Dev *23*, 419-432.

Nagy, L., and Schwabe, J. W. R. (2004). Mechanism of the nuclear receptor molecular switch. Trends in Biochemical Sciences *29*, 317-324.

Narasimhan, S. D., Yen, K., and Tissenbaum, H. A. (2009). Converging Pathways in Lifespan Regulation. Current Biology *19*, R657-R666.

Nuclear Receptor Committee (1999). A unified nomenclature system for the nuclear receptor superfamily. Cell *97*, 161-163.

Ogg, S., Paradis, S., Gottlieb, S., Patterson, G. I., Lee, L., Tissenbaum, H. A., and Ruvkun, G. (1997). The Fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in *C. elegans*. Nature *389*, 994-999.

Oh, S. W., Mukhopadhyay, A., Svrzikapa, N., Jiang, F., Davis, R. J., and Tissenbaum, H. A. (2005). JNK regulates lifespan in Caenorhabditis *elegans* by modulating nuclear translocation of forkhead transcription factor/DAF-16. Proceedings of the National Academy of Sciences of the United States of America *102*, 4494-4499.

Ow, M. C., Martinez, N. J., Olsen, P. H., Silverman, H. S., Barrasa, M. I., Conradt, B., Walhout, A. J. M., and Ambros, V. (2008). The FLYWCH transcription factors FLH-1, FLH-2, and FLH-3 repress embryonic expression of microRNA genes in *C. elegans*. Genes Dev *22*, 2520-2534.

Palanker, L., Tennessen, J. M., Lam, G., and Thummel, C. S. (2009). Drosophila HNF4 regulates lipid mobilization and beta-oxidation. Cell Metab *9*, 228-239.

Pan, D., Fujimoto, M., Lopes, A., and Wang, Y. (2009). Twist-1 is a PPARdelta-inducible, negative-feedback regulator of PGC-1alpha in brown fat metabolism. Cell *137*, 73-86.

Parter, M., Kashtan, N., and Alon, U. (2007). Environmental variability and modularity of bacterial metabolic networks. BMC Evolutionary Biology *7*, 169.

Piano, F. (2006). *C. elegans* network biology: a beginning. In WormBook Available at: http://www.wormbook.org/chapters/www_networkbio/networkbio.html [Accessed July 11, 2010].

Rabbitts, B. M., Ciotti, M. K., Miller, N. E., Kramer, M., Lawrenson, A. L., Levitte, S., Kremer, S., Kwan, E., Weis, A. M., and Hermann, G. J. (2008). glo-3, a novel Caenorhabditis *elegans* gene, is required for lysosome-related organelle biogenesis. Genetics *180*, 857-871.

Rando, O. J., and Chang, H. Y. (2009). Genome-wide views of chromatin structure. Annu. Rev. Biochem *78*, 245-271.

Rastinejad, F., Perlmann, T., Evans, R. M., and Sigler, P. B. (1995). Structural determinants of nuclear receptor assembly on DNA direct repeats. Nature *375*, 203-211.

Ravasz, E., Somera, A. L., Mongru, D. A., Oltvai, Z. N., and Barabási, A. L. (2002). Hierarchical organization of modularity in metabolic networks. Science *297*, 1551-1555.

Reece-Hoyes, J. S., Deplancke, B., Shingles, J., Grove, C. A., Hope, I. A., and Walhout, A. J. M. (2005). A compendium of Caenorhabditis *elegans* regulatory transcription factors: a resource for mapping transcription regulatory networks. Genome Biol *6*, R110.

Reece-Hoyes, J. S., Shingles, J., Dupuy, D., Grove, C. A., Walhout, A. J. M., Vidal, M., and Hope, I. A. (2007). Insight into transcription factor gene duplication from Caenorhabditis *elegans* Promoterome-driven expression patterns. BMC Genomics *8*, 27.

Reece-Hoyes, J. S., Deplancke, B., Barrasa, M. I., Hatzold, J., Smit, R. B., Arda, H. E., Pope, P. A., Gaudet, J., Conradt, B., and Walhout, A. J. M. (2009). The *C. elegans* Snail homolog CES-1 can activate gene expression in vivo and share targets with bHLH transcription factors. Nucl. Acids Res. *37*, 3689-3698.

Resendis-Antonio, O., Freyre-González, J. A., Menchaca-Méndez, R., Gutiérrez-Ríos, R. M., Martínez-Antonio, A., Avila-Sánchez, C., and Collado-Vides, J. (2005). Modular analysis of the transcriptional regulatory network of E. coli. Trends Genet *21*, 16-20.

Ririe, T. O., Fernandes, J. S., and Sternberg, P. W. (2008). The Caenorhabditis *elegans* vulva: a post-embryonic gene regulatory network controlling organogenesis. Proc. Natl. Acad. Sci. U.S.A *105*, 20095-20099.

Robinson-Rechavi, M., Maina, C. V., Gissendanner, C. R., Laudet, V., and Sluder, A. (2005). Explosive lineage-specific expansion of the orphan nuclear receptor HNF4 in nematodes. J. Mol. Evol *60*, 577-586.

Rosenfeld, M. G., Lunyak, V. V., and Glass, C. K. (2006). Sensors and signals: a coactivator/corepressor/epigenetic code for integrating

signal-dependent programs of transcriptional response. Genes & Development *20*, 1405-1428.

Rual, J., Ceron, J., Koreth, J., Hao, T., Nicot, A., Hirozane-Kishikawa, T., Vandenhaute, J., Orkin, S. H., Hill, D. E., van den Heuvel, S., et al. (2004). Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library. Genome Res *14*, 2162-2168.

Salgado, H., Gama-Castro, S., Peralta-Gil, M., Diaz-Peredo, E., Sanchez-Solano, F., Santos-Zavaleta, A., Martinez-Flores, I., Jimenez-Jacinto, V., Bonavides-Martinez, C., Segura-Salazar, J., et al. (2006). RegulonDB (version 5.0): Escherichia coli K-12 transcriptional regulatory network, operon organization, and growth conditions. Nucl. Acids Res. *34*, D394-397.

Sandmann, T., Girardot, C., Brehme, M., Tongprasit, W., Stolc, V., and Furlong, E. E. (2007). A core transcriptional network for early mesoderm development in Drosophila melanogaster. Genes & Development *21*, 436-449.

Schroeder, L. K., Kremer, S., Kramer, M. J., Currie, E., Kwan, E., Watts, J. L., Lawrenson, A. L., and Hermann, G. J. (2007). Function of the *Caenorhabditis elegans* ABC transporter PGP-2 in the biogenesis of a lysosome-related fat storage organelle. Mol. Biol. Cell *18*, 995-1008.

Scott, A., and Ghedin, E. (2009). The genome of Brugia malayi – all worms are not created equal. Parasitol Int *58*, 6-11.

Segal, E., Shapira, M., Regev, A., Pe'er, D., Botstein, D., Koller, D., and Friedman, N. (2003). Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. Nat. Genet *34*, 166-176.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res *13*, 2498-2504.

Shostak, Y., Van Gilst, M. R., Antebi, A., and Yamamoto, K. R. (2004). Identification of *C. elegans* DAF-12-binding sites, response elements, and target genes. Genes Dev *18*, 2529-2544.

Sladek, F. M., Zhong, W. M., Lai, E., and Darnell, J. E. (1990). Liver-enriched transcription factor HNF-4 is a novel member of the steroid hormone receptor superfamily. Genes Dev *4*, 2353-2365.

Sluder, A. E., and Maina, C. V. (2001). Nuclear receptors in nematodes: themes and variations. Trends Genet *17*, 206-213.

Sluder, A. E., Mathews, S. W., Hough, D., Yin, V. P., and Maina, C. V. (1999). The nuclear receptor superfamily has undergone extensive proliferation and diversification in nematodes. Genome Res *9*, 103-120.

Sonoda, J., Pei, L., and Evans, R. M. (2008). Nuclear receptors: decoding metabolic disease. FEBS Lett *582*, 2-9.

Soukas, A. A., Kane, E. A., Carr, C. E., Melo, J. A., and Ruvkun, G. (2009). Rictor/TORC2 regulates fat metabolism, feeding, growth, and life span in *Caenorhabditis elegans*. Genes Dev *23*, 496-511.

Spieth, J., and Lawson, D. (2006). Overview of gene structure. WormBook, 1-10.

Srinivasan, S., Sadegh, L., Elle, I. C., Christensen, A. G. L., Faergeman, N. J., and Ashrafi, K. (2008). Serotonin regulates *C. elegans* fat and feeding through independent molecular mechanisms. Cell Metab *7*, 533-544.

St Johnston, D., and Nüsslein-Volhard, C. (1992). The origin of pattern and polarity in the Drosophila embryo. Cell *68*, 201-219.

van Steensel, B., Delrow, J., and Henikoff, S. (2001). Chromatin profiling using targeted DNA adenine methyltransferase. Nat. Genet *27*, 304-308.

Takahashi, K., Tanabe, K., Ohnuki, M., Narita, M., Ichisaka, T., Tomoda, K., and Yamanaka, S. (2007). Induction of Pluripotent Stem Cells from Adult Human Fibroblasts by Defined Factors. Cell *131*, 861-872.

Taubert, S., Van Gilst, M. R., Hansen, M., and Yamamoto, K. R. (2006). A Mediator subunit, MDT-15, integrates regulation of fatty acid metabolism by NHR-49-dependent and -independent pathways in *C. elegans*. Genes Dev *20*, 1137-1149.

Taubert, S., Ward, J. D., and Yamamoto, K. R. (2010). Nuclear hormone receptors in nematodes: Evolution and function. Molecular and Cellular Endocrinology *In Press, Uncorrected Proof*. Available at: http://www.sciencedirect.com.ezproxy.umassmed.edu/science/article/B6T3G-5004R48-1/2/5bae9a8c2755cfa9d662a6c3674d68c5 [Accessed June 30, 2010].

Van Gilst, M., Gissendanner, C. R., and Sluder, A. E. (2002). Diversity and function of orphan nuclear receptors in nematodes. Crit. Rev. Eukaryot. Gene Expr *12*, 65-88.

Van Gilst, M. R., Hadjivassiliou, H., Jolly, A., and Yamamoto, K. R. (2005a). Nuclear hormone receptor NHR-49 controls fat consumption and fatty acid composition in *C. elegans*. PLoS Biol *3*, e53.

Van Gilst, M. R., Hadjivassiliou, H., and Yamamoto, K. R. (2005b). A *Caenorhabditis elegans* nutrient response system partially dependent on nuclear receptor NHR-49. Proc. Natl. Acad. Sci. U.S.A *102*, 13496-13501.

Vaquerizas, J. M., Kummerfeld, S. K., Teichmann, S. A., and Luscombe, N. M. (2009). A census of human transcription factors: function, expression and evolution. Nat. Rev. Genet *10*, 252-263.

Vermeirssen, V., Barrasa, M. I., Hidalgo, C. A., Babon, J. A. B., Sequerra, R., Doucette-Stamm, L., Barabási, A., and Walhout, A. J. M. (2007a). Transcription factor modularity in a gene-centered *C. elegans* core neuronal protein-DNA interaction network. Genome Res *17*, 1061-1071.

Vermeirssen, V., Deplancke, B., Barrasa, M. I., Reece-Hoyes, J. S., Arda, H. E., Grove, C. A., Martinez, N. J., Sequerra, R., Doucette-Stamm, L., Brent, M. R., et al. (2007b). Matrix and Steiner-triple-system smart pooling assays for high-performance transcription regulatory network mapping. Nat. Methods *4*, 659-664.

Wagner, G. P., Pavlicev, M., and Cheverud, J. M. (2007). The road to modularity. Nat Rev Genet *8*, 921-931.

Walhout, A. J., Sordella, R., Lu, X., Hartley, J. L., Temple, G. F., Brasch, M. A., Thierry-Mieg, N., and Vidal, M. (2000a). Protein interaction mapping in *C. elegans* using proteins involved in vulval development. Science *287*, 116-122.

Walhout, A. J., Temple, G. F., Brasch, M. A., Hartley, J. L., Lorson, M. A., van den Heuvel, S., and Vidal, M. (2000b). GATEWAY recombinational cloning: application to the cloning of large numbers of open reading frames or ORFeomes. Meth. Enzymol *328*, 575-592.

Walhout, A. J., and Vidal, M. (2001). High-throughput yeast two-hybrid assays for large-scale protein interaction mapping. Methods *24*, 297-306.

Walhout, A. J. (2006). Unraveling transcription regulatory networks by protein–DNA and protein–protein interaction mapping. Genome Research *16*, 1445-1454.

Watt, A. J., Garrison, W. D., and Duncan, S. A. (2003). HNF4: a central regulator of hepatocyte differentiation and function. Hepatology *37*, 1249-1253.

Watts, J. L., and Browse, J. (2002). Genetic dissection of polyunsaturated fatty acid synthesis in *Caenorhabditis elegans*. Proceedings of the National Academy of Sciences of the United States of America *99*, 5854-5859.

Wisely, G. B., Miller, A. B., Davis, R. G., Thornquest, A. D., Johnson, R., Spitzer, T., Sefler, A., Shearer, B., Moore, J. T., Miller, A. B., et al. (2002). Hepatocyte Nuclear Factor 4 Is a Transcription Factor that Constitutively Binds Fatty Acids. Structure *10*, 1225-1234.

Wood, W. B. (1988). Introduction to *C. elegans* biology. In The nematode *C. elegans* (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press), pp. 1-16.

Yamagata, K., Furuta, H., Oda, N., Kaisaki, P. J., Menzel, S., Cox, N. J., Fajans, S. S., Signorini, S., Stoffel, M., and Bell, G. I. (1996). Mutations in the hepatocyte nuclear factor-4alpha gene in maturity-onset diabetes of the young (MODY1). Nature *384*, 458-460.

Yang, F., Vought, B. W., Satterlee, J. S., Walker, A. K., Jim Sun, Z., Watts, J. L., DeBeaumont, R., Saito, R. M., Hyberts, S. G., Yang, S., et al. (2006). An ARC/Mediator subunit required for SREBP control of cholesterol and lipid homeostasis. Nature *442*, 700-704.

Yu, H., Greenbaum, D., Xin Lu, H., Zhu, X., and Gerstein, M. (2004). Genomic analysis of essentiality within protein networks. Trends Genet *20*, 227-231.

Zeitlinger, J., Stark, A., Kellis, M., Hong, J., Nechaev, S., Adelman, K., Levine, M., and Young, R. A. (2007). RNA polymerase stalling at developmental control genes in the Drosophila melanogaster embryo. Nat. Genet *39*, 1512-1516.

Zhong, W., and Sternberg, P. W. (2006). Genome-Wide Prediction of *C. elegans* Genetic Interactions. Science *311*, 1481-1484.

Zilliacus, J., Wright, A. P., Carlstedt-Duke, J., and Gustafsson, J. A. (1995). Structural determinants of DNA-binding specificity by steroid receptors. Mol. Endocrinol *9*, 389-400.