Fall 12-7-2011

Transcriptional Targets of the REF-1 Family Proteins: HLH-25/ HLH-28/HLH-29

Kun Wang

TRANSCRIPTIONAL TARGETS OF THE REF-1 FAMILY PROTEINS: HLH-25/ HLH-28/HLH-29

by

KUN WANG

Under the Direction of Casonya Johnson

ABSTRACT

Notch signaling is important for development in *Caenorhabditis elegans* and the REF-1 family proteins, a set of the bHLH transcription factors, are the first targets of Notch signaling. Little is known about the molecular mechanisms employed by the REF-1 family to regulate development. In this project, I identified novel targets of three REF-1 family proteins, HLH-25/HLH-28/HLH-29, and determined which target genes are activated and which are repressed by the REF-1 proteins. These targets were identified by gene expression microarray and were functionally categorized by Gene Oncology analysis. A systems biology approach was performed to identify networks associated with those targets. In addition to the molecular genetics studies, I identified and better characterized the range of phenotypes induced by mutations in ref-1 family genes.

INDEX WORDS: REF-1 family, Transcription factors, Gene expression microarray

TRANSCRIPTIONAL TARGETS OF THE REF-1 FAMILY PROTEINS: HLH-29/ HLH-28/HLH-25

by

KUN WANG

A Thesis Submitted in Partial Fullfillment of the Requirement for the Degree of

Master of Science

in the College of Arts and Sciences

Georgia State University

2011

TRANSCRIPTIONAL TARGETS OF THE REF-1 FAMILY PROTEINS: HLH-29/ HLH-28/HLH-25

by

KUN WANG

Committee Chair: Walter (Bill) William Walthall

Committee: Casonya Matese Johnson

Parjit Kaur

Electronic Version Approved:

Office of Graduate Studies

College of Arts and Sciences

Georgia State University

December 2011

ACKNOWLEDGEMENTS

I am fortunate to have had my favorite professors on my thesis committee. First I must thank Dr. Casonya Johnson for her patient and thorough advice throughout my graduate studies. I would particularly like to thank the members of my committee, Dr. Walthall and Dr. Kaur, for their guidance and insights. Many thanks to members of the Johnson and Walthall labs, Chaquettea M Felton, Ana White, Huihui Zhu, Yanqing Zhou, Abir Ashfakur Rahman and Deep Shukla. You all gave me with the tools and support I needed at the right times and I wish you all great success.

I must thank my family including my parents, Yuqin Zhou and Jinhong Wang, my parents-in-law, Dr.Yang Xu, Esheng Zhou, and my wonderful husband, Yanqing Zhou. You give me with much support, comfort, and joy.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 INTRODUCTION

The Notch signaling pathway is a highly conserved cell signaling system which plays a prominent role in mediating cell-cell interactions during animal development. In mammals, Notch signaling regulates neuronal function and development (Gaiano et al., 2002; Del Monte et al., 2007), and stabilization of arterial endothelial fate and angiogenesis (Liu et al., 2003).

Members of the Hairy and Enhancer of Split (E(spl)) proteins, a group of bHLH transcription factors firstly described in *Drosophila*, are direct targets of Notch signaling (Jennings et al., 1994; Dawson et al., 1995; Alifragis et al., 1997; Giagtzoglou et al., 2003). Hairy/HES proteins are widely present in animals and serve as transcription factors during the developmental process. In addition to the well conserved HES/HAIRY proteins, animals also have other HES-like proteins, such as the HER and Hey proteins, which function similarly to HES but lack canonical HES-domains (Ninkovic at el., 2003; Leimeister et al., 2000). In the zebrafish midbrain, Her5 is found to be necessary for formation of the medial (most basal) part of intervening zone (MIZ) during neurogenesis, and inhibits expression of the proneural gene *ngn1*, in a dose-dependent manner (Ninkovic et al., 2003). In chick, cHey2 is expressed rhythmically across the chicken presomitic mesoderm (Leimeister et al., 2000) to control segmentation. In Xenopus, xHes2 acts in a Notch-dependent fashion to control neural cell fate decisions during retinogenesis and inhibits neuronal differentiation (Sölter at el., 2006).

From these studies, it is evident that Hairy/HES proteins, like most of the bHLH proteins, play significant roles in embryonic and post-embryonic development. Of the 42 bHLH genes in *C. elegans*, only six are considered to be potential HES-like proteins: REF-1, HLH-25, HLH-26, HLH-27, HLH-28 and HLH-29 (Neves et al., 2005), and are collectively considered to be the REF-1 family of proteins. First, these six proteins have homology within their bHLH domain to

the Hes/Hairy bHLH domain. Second, the REF-1 proteins are activated during Notch signaling events in *C.elegans* embryos. Third, these proteins are expressed in Notch activated cells during early development.

Structurally, the REF-1 family proteins are novel in that they each have two distinct bHLH domains (Alper et al., 2001). Each bHLH is homlogus to the bHLH domain of Hairy/E (spl)/Hes/Hey proteins, and overall, the REF-1 proteins are structurally similar (see Figure1.1). However, REF-1 family proteins lack the Orange and the C-terminal WRPW domains (Neves et al., 2005) found in canonical HES proteins. Instead, the C-terminal end of the six proteins contains a pentapeptide repeat sequence which is believed to serve similar functions to the WRPW domain (Neves et al., 2005). Of the six family members, the pentapeptide sequence of REF-1 protein is very similar to the WRPW domain – FRPWE. The other sequences are: LDIIN in HLH-25, IDIVG in HLH-26, VDISN in HLH-27, IDIIG in HLH-28 and IDIIG in HLH-29 (Figure 1.1).

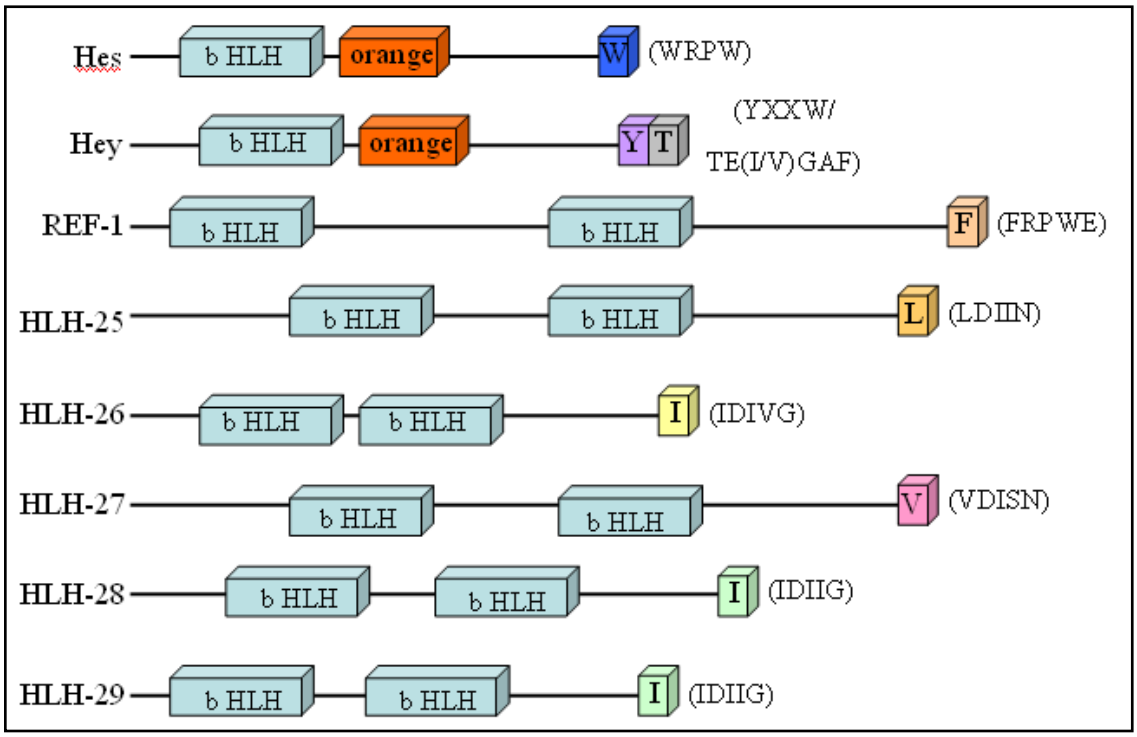

Figure 1.1. Domain Organization of Hes, Hey and REF-1 Family Proteins. This figure was adapted from Dawson at el.1995 and Neves et al. 2005. Hey proteins lack the WRPW tetrapeptide sequences found in HES proteins, and instead containing a related YRPW peptide or a further degenerated YXXW sequence. This tetrapeptide sequence is followed by a conserved TE(IV)GAF peptide. REF-1 family proteins lack the Orange domain. Our hypothesis is that the second basic helix-loop-helix domain can functionally substitute for the orange domain. Rather than WRPW in Hes or YXXW in Hey, REF-1 family proteins instead contain pentapeptide sequences that are similarly charged.

Mutations in the REF-1 family proteins support the hypothesis that these proteins play critical roles in development. REF-1 protein, the product of the gene *hlh-24*, regulates the expression of two Hox genes, *lin-39* and *mab-5*, whose products work together to control the cell fusion decisions during vulva development (Alper et al., 2001). In addition to affecting organ development, REF-1 acts in the neuronal lineage to suppress the generation of ectopic neurons from early neuronal or non-neuronal precursor cells via a Notch-independent pathway. For example, an embryonically generated serotinergic neuron type differentiates abnormally in animals with loss of function mutations in REF-1 (Lanjuin A et al., 2005).

The protein product of the *hlh-28* gene is identical to the product of the *hlh-29* gene (McMiller et al., 2006). So, while the expression of the two genes may differ in timing and location, the genes are likely to have identical targets and functions. Loss of HLH-28 and HLH-29 results in late embryonic lethality, yolk protein accumulation, everted vulva, and abnormal bordering behavior (McMiller et al., 2007). Both *hlh-28* and *hlh-29* are expressed in early embryogenesis in response to Notch signaling; *hlh-29* is expressed post-embryonically in both neuronal and non-neuronal tissues (McMiller et al., 2007). During early larval development, HLH-29 can rescue morphological phenotypes seen in REF-1 mutant animals, suggesting that these two proteins have overlapping functions (Neves et al., 2005). Recently, unpublished results from our laboratory suggest that HLH-29 is needed for proper regulation of ovulation and fertilization, further underscoring the importance of this protein family in *C. elegans* growth and development. The other REF-1 family proteins have not yet been genetically characterized, and complete loss of function alleles have not been isolated for all of the genes.

Like other bHLH proteins, the REF-1 family proteins are thought to bind to DNA as dimers. In other organisms, the HES proteins traditionally form heterodimers with a wide range of tran-

scription factors, including other bHLH proteins. A recent study suggests that, with the exception of HLH-26, the REF-1 family proteins do not heterodimerize with other bHLH proteins, but probably bind to DNA as homodimers (Grove et al., 2009). Using protein binding microarrays (PBM), Grove and others (2009) identified preferential DNA binding sequences for each of the REF-1 family proteins. HLH-26 was shown to heterodimerize with the Max-like protein, MDL-1/MXL-1, and to preferentially bind to the E box sequence CACGTG (Grove et al., 2009). The other REF-1 family proteins bind very selectively to non-cannonical E-box sequences. While most bHLH proteins bind stably to multiple variations of the canonical E-box (CACGTG), the REF-1 proteins are more selective in the sequences that they prefer. Nevertheless, HLH-25 can recognize five different E-box or E-box-like sequences. In contrast, HLH-26 and REF-1 bind exclusively to CACGTG (Grove et al., 2009). Using a bioinformatics approach, putative target genes were identified for each REF-1 family protein based on the proximity of the specific DNA binding sequence to the translation initiation site of genes within the *C. elegans* genome. These target genes were then grouped by function using gene ontology (GO) analysis. Based on this analysis, HLH-25 is the only REF-1 family protein associated with "cell division" (Grove et al., 2009). Additionally, HLH-25 is connected to nine other GO terms, an unusually high number for bHLH proteins in general. Together, these studies suggest that the REF-1 protein family is critical for development in *C.elegans*; the GO analysis from the microarray experiment in my project also suggest that both HLH-25-regulated and HLH-29-regulated genes highly contribute to development processes.

Based on the structural and functional comparisons between the REF-1 family in *C. elegans* and Hairy/HES in other animals, it is likely that the REF-1family plays an important role in development in *C. elegans*. My objective is to understand how HLH-25 and HLH-29 function to

regulate *C. elegans* development. One way to do this is by identifying transcriptional targets of HLH-25 and HLH-29 and then using a systems biology approach to identify networks associated with those targets. Here I describe the results of this approach using gene expression microarray.

2 MATERIALS AND METHODS

2.1 *C. elegans* Growth and Culture Conditions

hlh-25 (*ok1710*) mutant animals were acquired from the Caenorhabditis Genomics Center (CGC), *hlh-29* (*tm284*) mutant animals were acquired from Tokyo Women's Medical College (Japan) and were maintained at 20°C on NGM agar plates seeded with *Escherichia coli* strain OP50 as a food source as previously described (Sulston and Hodgkin, 1988).

2.2 RNAi Feeding

RNAi clones were fed *C. elegans* as described previously (Kamath and Ahringer 2003). All animals were synchronized to L1 stage by hypochlorite treatement of gravid adults (Kamath and Ahringer 2003) and were fed control bacteria or RNAi-producing bacteria until adult stage. All strains were maintained at 20°C on NGM as previously described. The strain tm284 was fed bacteria producing the *hlh-28/hlh-29* dsRNA. The strains N_2 and *ok1710* both were fed bacteria with control clone producing dsRNA for an unrelated gene.

2.3 Total RNA Isolation

Total RNA was extracted essentially as described by Kostrouchova et al., 2001. The worms in L4 or adult stage were collected by washing several times with water and pelleting by centrifugation for 1 min at 2,000 rpm. All samples were frozen at −80°C. Each pellet was resuspended in 0.5 ml of buffer containing 0.5% SDS, 5% 2-mercaptoethanol, 10 mM EDTA, 10 mM Tris· HCl (pH 7.5), and 0.5 mg/ml proteinase K. The samples were incubated 1 hr at 55°C and extracted with 1 vol of phenol/chloroform (1:1). The RNA was precipitated with ethanol at −80°C overnight, collected by centrifugation, and resuspended in nuclease-free water. The samples were

treated with proteinase K with concentration 50µg/ml for 1 hr at 50°C, extracted twice with phenol/chloroform, and precipitated in ethanol. The extracted RNA was dissolved in nuclease-free water, and its concentration and purity were determined from absorbance measurements at 260 and 280 nm using a spectrophotometer

2.4 cDNA Synthesis

Total RNA from three biological replicates from each strain was used to make cDNA (Severance 2010). cDNA was synthesized using the high capacity cDNA reverse transcription kit (Applied Biosciences catalog #4368814) as directed by the manufacturer.

2.5 Gene Expression Microarray

Gene expression microarray was performed by GSU DNA/Protein Core Facility. Global gene expression in synchronized populations of *ok1710* or *tm284* animals was compared to the expression of N2 (wild-type) animals using GeneChip *C. elegans* Genome Array (Affymetrix). Data collection was carried out using GCOS 1.4 software (Affymetrix).

2.6 Microarray Data Analysis

2.6.1 Normalization and Quality Controls

GeneSpring *GX 11* Software (Aglient, Palo Alto,CA) was used to carry out data-analysis. For the probe intensity values generated by the Affymetrix scanner, Robust Multichip Average (RMA)-algorithm was used to normalize the collected data. The quality controls on samples and on probe sets were performed stepwise to detect the outlying samples and the poor probe sets. For samples detection, the Principal Components Analysis (PCA) score plot and hybridization

controls plot were applied for detecting. The poor quality samples present the deviation from the expected intensity profile of these controls in hybridization controls plot, while, in PCA plot, samples under the same experimental condition should present more similar to each other than to samples under a different condition. For probe sets detection, the aim was to delete low-intensity signals of genes that are not expressed (http://genespring.com).

2.6.2 Significance Analysis

Based on the above process, the good quality samples and the probe sets with reliable intensity measurements were used for statistical analysis. The T test was performed to find the candidates for differential expression, and genes with significant signal level between different conditions ($p<0.05$) were collected. In addition, fold change analysis were performed on the genes with significant expression. In this experiment, >2-fold-change were chosen as the significant level.

2.7 Functional Analysis

2.7.1 Gene Ontology (GO) analysis

The Gene Ontology (GO) is a collection of controlled ontologies describing the functions of a gene product in any species (http://www.geneontology.org/GO.doc.shtml). There are three sets of ontologies: cellular component (like “organelle membrane”), molecular function (like “DNA binding”) and biological process (like “embryonic development”). *Genespring* GX 11 software was used for GO analysis. Lists of HLH-25-, HLH-29- and control- regulated genes are uploaded and calculation of GO biological process annotations using function annotation clustering was performed using the GO-chart function. Significance cut-off was set at $P < 0.05$ (http://genespring.com).

2.7.2 Pathway Analysis

To further and better understand the gene functions, pathway analysis was performed on the gene list which was derived from significance and fold-change analysis in the microarray experiment by using *Genespring* GX11 software. The interactions in the database are derived from published literature using Natural Language Processing (NLP) algorithm. Relation score was calculated to indicate a confidence matrix on the quality of relations and had a scale of 1-9, where 9 is the best and 1 is the weakest. In pathway analysis, "connectivity" is another term used to describe how well the gene is connected to other gene. There are two types of connectivity: the "local connectivity" of a gene refers to the number of other connected genes while "global connectivity" of a gene refers to how many relations the gene participates in (http://genespring.com).

2.8 Life Span Assay

Life span assay was performed as previously described (Larsen et al. 2002). Assays were conducted at 16°C and 25°C. Animals were synchronized by hypochlorite treatment, fed RNAi until L3 stage and then 150 animals were already eating the appropriate RNAi from hatching, were transferred to fresh RNAi NGM plates. There were two replicates for each strain. During the egg laying period, these worms were transferred every two days to a new fresh NGM plates, seeded with RNAi-producing bacteria. The numbers of survival, dead and missing worms were counted each time. The Graphpad Software Package was used for statistical analysis and to calculate means and percentiles. In all cases p-values were calculated using the Mantel-Cox (logrank) test.

3 RESULTS

3.1 DNA Microarray Data Analysis

Statistical analysis of the microarray data was initially performed using Genespring GX11 software. For *hlh-25* samples, 17981 out of the 22627 probe sets passed the probe sets filter and were used to perform the subsequent t-test and fold change analysis. We set the minimum criteria at 2-fold change and a p-value<0.05 after the t-test. We observed 634 regulated genes which met the criteria of which 510 were up-regulated and 124 were down-regulated. These genes are listed in supplemental Table S1. For the *hlh-29* samples, 284 genes met the criteria, of which 250 were up-regulated and 34 were down-regulated (Table S2).

To find the genes that were regulated by both HLH-25 and HLH-29, an overlapping test was performed on the significantly expressed genes. As shown in Figure 3.1, 80 affymetrix tags, representing 71 genes are regulated by both HLH-25 and HLH-29. As shown in Table 3.1, 48 genes were up-regulated, and 6 genes were down-regulated in both, and 17 genes were regulated in different manner.

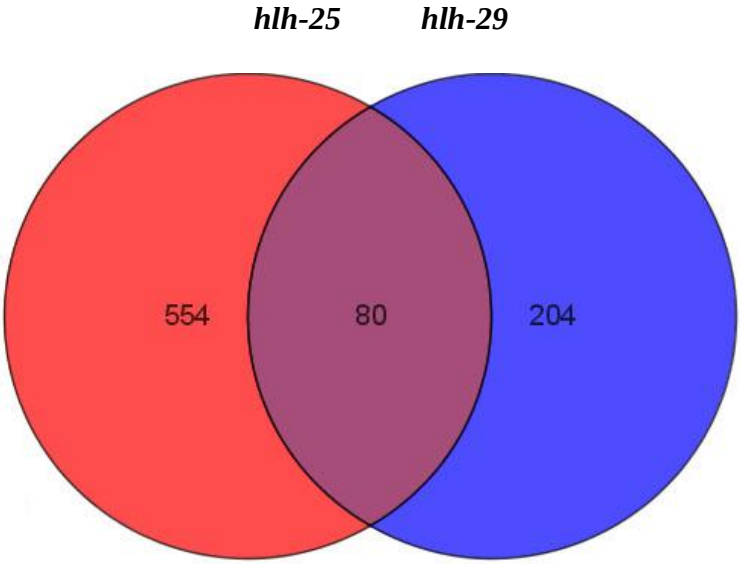

Figure 3.1 Venn-diagram of genes whose expression changed at least 2.0 fold in *hlh-25* mutants and *hlh-2*9 mutants compared to N2.

Table 3.1 Genes Changed >2 Fold In *hlh-25* and *hlh-29* Mutant Animals

Gene Symbol	Fold Change of HLH-25	Regulation of HLH-25	Fold Change of HLH-29	Regulation of HLH-29
anc-1	2.0620477	up	2.1114826	up
atp-3	2.2059124	up	2.0204434	up
C01B10.6	2.395367	up	2.029442	up
C04F12.7	4.2752776	up	3.4203176	up
C10G11.9	3.7218602	up	4.643113	up
C15C6.2	2.876846	up	2.6460059	up
C15C8.3	2.1991796	up	2.318762	up
C36A4.9	2.2960815	up	2.6848862	up
C39H7.1	2.4883223	up	2.1495948	up
C53B7.3	3.3917544	up	4.1526747	up
clec-60	2.1613488	up	2.4010558	up
cpg-2	2.4668424	up	2.1163735	up
egl-21	2.223056	up	2.0198085	up
emb-9	3.438133	up	2.3422027	up
F01G4.6	2.7248034	up	2.6292844	up
F29G6.3	3.6033294	up	2.0007153	up
F36H12.8	2.3870664	up	2.305656	up
F44E5.4	2.3633173	up	2.457278	up
F45D11.14	2.363508	up	2.8867633	up
F55H12.4	2.3432472	up	2.4337006	up
fat-2	2.0514073	up	3.5710244	up
his-48	2.1418402	up	3.017254	up
his-62	2.0827363	up	2.3219244	up
K01D12.15	2.1774943	up	2.0681512	up
K06A5.2	2.8447423	up	2.3556824	up
lin-37	3.0031235	up	2.135986	up
nurf-1	2.2031708	up	2.269614	up
pab-2	2.0896707	up	2.3890529	up
plp-1	2.5941017	up	2.147615	up
pos-1	3.7705986	up	2.3749654	up
ppn-1	3.2876306	up	2.7322247	up
ret-1	2.96265	up	2.0428646	up
sca-1	3.4480622	up	2.6404407	up
T08G11.1	2.7808566	up	2.113738	up
T23G11.1	2.4043186	up	2.343141	up
tag-18	2.3739333	up	2.289861	up
ttr-27	2.6074536	up	2.113668	up
ttr-44	2.1621397	up	2.936464	up
unc-15	4.2877297	up	2.723157	up

Table 3.1 Continued
Genes Changed >2 Fold In *hlh-25* and *hlh-29* Mutant Animals

Gene Symbol	Fold Change of HLH-25	Regulation of HLH-25	Fold Change of HLH-29	Regulation of HLH-29
W02B12.12	2.1325982	up	2.4280045	up
YER141W	2.166064	up	2.1664827	up
Y69E1A.2	2.7324333	up	2.05632	up
ZK829.4	5.580293	up	2.51748	up
ZK858.2	2.8894699	up	2.2106056	up
ZK616.6	3.1095974	up	2.0016468	up
xbp-1	2.29453	up	2.2447314	up
ZK484.1	2.5698814	up	2.6460128	up
mai-1	8.631056	up	3.2633963	up
T25C12.3	2.9362626	up	8.889503	down
F35E12.5	2.8423295	up	2.3189948	down
F59A2.5	2.2631724	down	2.3684695	down
ilys-2	5.4114156	down	2.030791	down
lys-10	7.699708	down	6.7607007	down
msh-5	3.2695198	down	2.3055043	down
sel-5	4.1654544	down	2.354605	down
Y39B6A.1	2.9412065	down	3.4929483	down
acs-2	3.2818496	down	6.4914927	up
cnc-7	3.2016509	down	2.4217145	up
F09F7.6	5.6009636	down	3.174787	up
F18E3.7	2.0175085	down	2.169345	up
F21C10.10	2.014651	down	2.6284351	up
F46A8.7	6.972377	down	2.0406806	up
fipr-23	2.1018212	down	4.9249797	up
fmo-2	3.0290232	down	2.2133856	up
gei-7	4.875362	down	2.774993	up
lea-1	2.1253364	down	4.3557854	up
mtl-1	17.338673	down	3.573709	up
T12D8.5	9.470627	down	3.8493972	up
T16G1.4	2.1095414	down	2.1513166	up
ZC395.5	3.002598	down	2.4850364	up
Y102A5C.6	2.8645706	down	2.5697527	up

3.2 Gene Ontology Analysis

To determine the functional distribution of the putative target genes, GO analysis was applied to genes affected >2.0-fold using the *Genespring* GX11 software. Based on the phenotypes seen in the REF-1 family mutants, and based on the genetic link between the REF-1 family proteins and the Notch signaling pathway, we made the hypothesis that HLH-25 and HLH-29 may play an important role in development events in *C. elegans*. In GO analysis, genes functions are assigned to one of three domains: Cellular Component, Molecular Function and Biological Process. Of those three domains, cellular events that are related to embryonic and post-embryonic development are included under the Biological Process domain; hence, I performed my annotation using the Biological Process domain. The summaries of GO annotation are shown in Tables 3.2 and 3.3 for HLH-25 and HLH-29, respectively. The full list of the GO annotations associated with HLH-25 and HLH-29 are listed in the Tables S3 and S4, respectively.

For HLH-25, 634 genes were assigned to GO terms with some overlapping annotation. The genes were most commonly found in the categories, in order from most significant to least: regulation of growth, embryonic development ending in birth or egg hatching and post-embryonic development. Embryonic development ending in birth or egg hatching has the highest percentage compared with other terms. Four GO terms related to growth also have highly percentages (Table 3.2). These results supported my hypothesis that HLH-25 plays an important role in embryonic and nematode larval development. For HLH-29, 284 genes were assigned to GO terms with some overlapping annotation. These genes were grouped in the categories of: aging, oxidation reduction, and positive regulation of growth (Table 3.3). These result also supported my hypothesis, though they suggest that HLH-29 functions more in post-embryonic development, during mid and late larval stages. Interestingly, only a small percentage of the genes affected by loss of

hlh-29 were associated with terms directly related to embryonic development (see supplemental table S4).

Additionally, to identify common functional distribution between HLH-25 and HLH-29, the GO analysis was applied to genes regulated by both HLH-25 and HLH-29 (Table 3.4). Collectively, these genes were assigned to the categories: embryonic development ending in birth or egg hatching, regulation of growth rate, and growth.

Table 3.2 Gene Ontology Biological Process for HLH-25

GO term(biological process)	Count	%	P-value
Regulation of growth	136	22.8	2.9E-09
Embryonic development ending in birth or egg hatching	168	28.1	3.8E-09
Nematode larval development	122	20.4	3.2E-08
Positive regulation of growth rate	120	20.1	3.4E-08
Positive regulation of growth	129	21.6	3.9E-08
Post-embryonic development	122	20.4	6E-08
Body morphogenesis	53	8.9	5.2E-07
Growth	96	16.1	0.0000011

Terms associated with genes list from *hlh-25* mutant animals (FC>2.0.). The count indicates the number of observations from the input of 634 genes. % count indicates the percentage of genes in the input entity list which have that GO term. P-values are the probabilities of obtaining the specified GO accession number of genes in the each GO term. Terms are listed in decreasing order of significance (p-value).

Table 3.3 Gene Ontology Biological Process for HLH-29

GO term(biological process)	Count	%	P-value
Aging	14	5.5	4.70E-05
Multicellular organismal aging	14	5.5	4.70E-05
Determination of adult life span	14	5.5	4.70E-05
Oxidation reduction	17	6.6	1.20E-04
Positive regulation of growth rate	34	13.3	4.80E-02
Regulation of growth rate	34	13.3	4.90E-02
Positive regulation of growth	36	14.1	7.00E-02

Terms associated with genes list from hlh-29 mutant animals (FC>2.0.). The count indicates the number of observations from the input of 284 genes. % count indicates the percentage of genes in the input entity list which have that GO term. P-values are the probabilities of obtaining the specified GO accession number of genes in the each GO term. Terms are listed in decreasing order of significance (p-value).

Table 3.4 Gene Ontology Biological Process for Genes Affected by HLH-25 and HLH-29

GO Term(biological process)	Count	%	P-value
Regulation of growth rate	16	22.5	8.40E-03
Positive regulation of growth	16	22.5	2.10E-02
Aging	5	7	2.70E-02
Determination of adult life span	5	7	2.70E-02
Cytokinesis after mitosis	2	2.8	3.30E-02
Embryonic development ending in birth or egg hatching	19	26.8	4.00E-02
Cytokinesis during cell cycle	2	2.8	5.10E-02
Growth	11	15.5	9.60E-02
Macromolecular complex assembly	3	4.2	9.60E-02

Terms associated with genes whose expression is affected in both hlh-29 and hlh-25 mutant animals. The count indicates the number of observations from the input of 71 genes. % count indicates the percentage of genes in the input entity list which have that GO term. P-values are the probabilities of obtaining the specified GO accession number of genes in the each GO term. Terms are listed in decreasing order of significance (p-value).

3.3 HLH-25 and HLH-29 Functional Networks in Development

We identified HLH-25 and HLH-29 target genes that work in functional pathways, or networks, in an effort to further understand the biological role of the two REF-1 proteins. Using the software analysis program already available through Genespring, we only considered protein functions in complex networks which can include binding interaction networks, transcription regulation networks, or translation regulations networks. A summary of this analysis is represented in Figure 3.3, with both "binders" and "regulators" indicated in blue and yellow, respectively. In this study, "binders" refers to any target gene product that is associated with two or other genes in a given pathway, through either direct or indirect protein-protein interactions. The term "regulators" refers to genes that function upstream of other genes, and that are required for proper functioning of the downstream gene, which would be considered the "network target" gene. For HLH-25, 53 out of 634 putative targets are involved in binding networks and regulation networks. For HLH-29, 42 out of 254 genes are identified to be involved in binding networks and regulation networks (Figure 3.3). Although 42 of the HLH-29 targets were found to be involved in binding networks and regulation networks, very few of the individual networks identified involved more than two HLH-29 targets. It was difficult, from the information gathered to date, to determine how and why those 42 genes would be linked together by one transcriptional regulator. Therefore, the analysis presented here focuses primarily on HLH-25targets, and future work in the lab will more carefully assess the HLH-29 targets.

3.3.1 HLH-25 Regulates Cell Division in *C. elegans*

Increased expression of three genes that mediate the establishment of embryonic polarity, *mex-5, mex-6,* and *pos-1*(2.1- to 3.8- fold) suggests that HLH-25 functions in controlling asymmetric cell division during embryo development (Figure 3.3.a). *mex-5* and *mex-6* encode two

CCCH-finger proteins that function to establish soma/germline asymmetry in the early embryo (Schubert et al. 2000). POS-1, another CCCH-finger protein, is required for fate specification of germ cells, intestine, pharynx and hypodermis during embryogenesis (Tabara et al. 1999). In the establishment of the germ line, soma/germline asymmetry, all three of these CCCH-finger proteins interact with ZIF-1, the SOCS-box protein, to degrade germ plasm proteins in somatic blastomeres (Derenzo et al.2003). We observed increased expression of these three genes in *hlh-25* mutants which suggest that HLH-25 normally represses the asymmetric cell division.

In addition to controlling genes needed for the segregation of cell fate determinants during cell division, HLH-25 affects the expression of genes that function in other cell division related events in the embryo (Figure 3.3.b). *zyg-9* encodes a microtubule-association protein (MAP) that functions to control pronuclear migration and spindle elongation in the one-cell-stage embryo (Bellanger and Gönczy 2003). Expression of this gene increased 2.7-fold in *hlh-25* mutants. The *cks-1*gene, which increases expression 3.3-fold in *hlh-29* mutants (Figure 3.3.b), encodes a highly conserved cell cycle regulatory protein and plays a critical role in endoderm specification and the spindle orientation (Polinko et al. 1997). The genes *ima-2* and *ran-1* encode proteins that are essential for chromosome segregation and mitotic spindle formation, respectively. IMA-2 in particular, is required during embryonic mitosis for nuclear reassembly (Geles et al. 2002; Srayko et al. 2005). Both of these genes are up-regulated more than 2-fold in *hlh-25* mutant animals (Figure 3.3.b). These results suggest that the expression of activated *hlh-25* may inhibit cell division by regulation these four genes.

3.3.2 HLH-25 Participates in the Wnt Signaling Pathway

The Wnt signaling pathway also plays a crucial role during embryonic and post-embryonic development. CGH-1, a putative DEAD-box RNA helicase (Walhout et al.2002), and C05C10.5,

a novel protein, both bind to one of the two *C. elegans* Dishevelled homologs (Walhout et al.2002), DSH-2, which functions via Wnt signaling to regulate cell fate specification and asymmetric cell division (Walhoutetal and et al, 2002; Kyla and et al, 2009). Likewise, the gene products of *tbb-1* and *lir-1* both bind to LIT-1, the *C.elegans* homolog of the Nemo protein, which acts via Wnt signaling to direct axis formation and embryonic polarity (Herman, 2001). Loss of HLH-25 causes increased expression of CGH-1, C05C10.5, TBB-1 and LIR-1(Figure 3.3.c). Taken together, our data suggest that HLH-25 functions embryonic cells to control devel-opmental events such as cell fate specification, mitosis, and asymmetric cell division.

3.3.3 HLH-25 and HLH-29 May Depress Notch Signaling via POS-1

As mentioned above, loss of HLH-25 increases *pos-1* expression by 3.8-fold (Figure 3.3.d). Interesting, POS-1 expression is also up-regulated in *hlh-29* mutant animals. POS-1 is essential for translational repression of GLP-1, one of the two Notch receptor proteins found in *C.elegans* (Farley et al.2010). This result is particularly interesting for two reasons. First, it suggests that HLH-25 and HLH-29 may enhance Notch signaling by derepressing the Notch receptor. Second, negative feedback regulation of Notch signaling by HES proteins has been previously reported (King and et al, 2006), further underscoring the possibility that the REF-1 proteins are functional HES-like proteins in *C.elegans*.

3.3.4 HLH-25 and HLH-29 Target Genes Associated with Longevity

In *C. elegan*s, numerous genes have been identified that directly or indirectly regulate lon-gevity most often by using the aging phenotype as a marker for extension of life span. Mutations that can extend or shorten the life-span of *C. elegans* have provided insight into molecular mech-anisms underlying aging and longevity. Longevity is affected by many pathways. For instance, DAF-2, like its mammalian homolog insulin/insulin-like growth factor 1(IGF1), functions in an

insulin-like signaling pathways which is recognized as a universal regulator of longevity (Pierce et al. 2001), and negatively regulates the FOXO transcriptional factor DAF-16 through a conserved PI-3 kinase/PDK/AKT, SGK cascade (Kenyon, 2005). The extension of life span was found to be dependent on wild-type function of DAF-16(Lin et al.1997).

Although our microarray results show that genes affected by both HLH-25 and HLH-29 are enriched in the GO term, "determinations of life span" (see Table 3.4). We were unable to find networks that clearly show the relationship between life span and these two genes. Interestingly, HLH-25 regulates the expression of AKT-2 (Figure 3.3.e), a serine/thereonine kinase that is a functional suppressor of DAF-16 (Quevedo et al.2007). It is also noteworthy that the DAF-16 target, SOD-3, is down-regulated in HLH-29 mutants (Figure 3.3.e). The SOD-3 gene encodes superoxide dismutase, and is also regulated by the insulin growth factor receptor gene, daf-2 (Vanfleteren et al. 1995). Together, these data provide an interesting molecular link between the REF-1 family and life span.

3.4 Life-Span Assay

We used the life span assay to assess potential differences in aging between wild-type animals eating control RNAi and *hlh-25* or *hlh-29* mutant animals. As shown in Figure 3.4, life-span is not affected in HLH-25 or HLH-29 animals when grown at 16°C. Unfortunately, the results at 25°C were conflicting and need to be repeated with a larger number of animals and probably using an incubator that has a tighter temperature regulator at 25°C. We repeated our assays twice at 25°C; in one assay, the results were the same as at 16°C. In other assay, loss of *hlh-25* increased lifespan while loss of *hlh-29* decreased life span. We do not show the results at 25°C

here because that assay was done by a different member of the lab. We believe these conflicting results may be due to the influence of other redundant genes, in regulating life span.

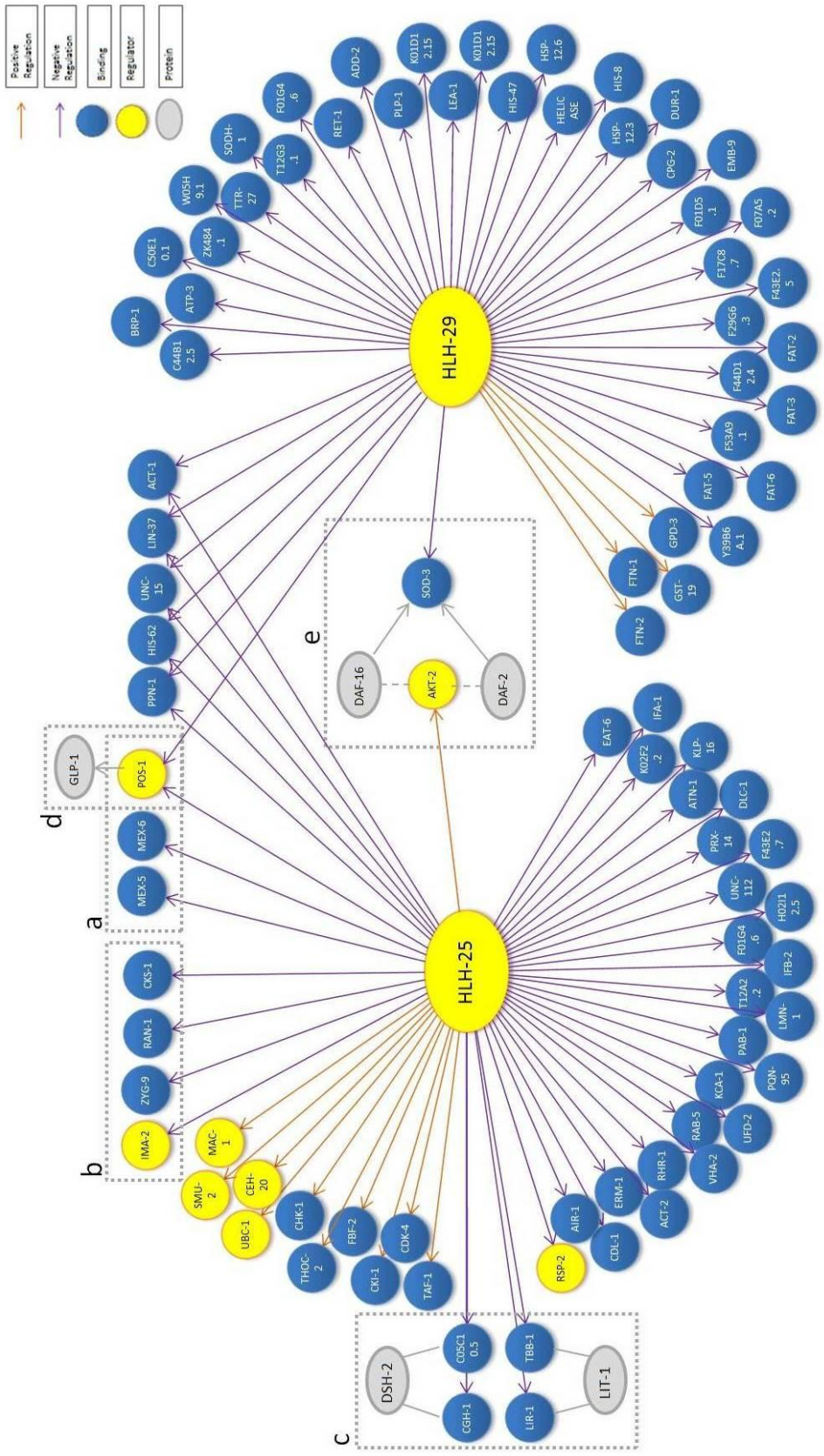

Figure 3.3 The graphical representation of binding and regulation networks around HLH-25 and HLH-29.

Figure 3.3 Graphical Representation of Binding and Regulation Networks Around HLH-25 and HLH-29. Blue cycles represent the "binders" - proteins and gene products that associate with at least two other genes in a given network. Binders represent multiple protein classes; examples include kinases, phosphatases, and ligand receptor proteins. Yellow cycles represent "regulators" - proteins that act upstream of a given gene in a regulatory role, most often as transcription factors. Grey cycles represent the proteins which were not identified in the gene expression array as targets of HLH-25 or HLH-29. Red and purple arrows represent positive and negative regulation, respectively, by HLH-25 or HLH-29. Five important biological processes or pathways described further in the text are designated as parts: a) and b) pathways known to be associated with cell division; c) Wnt Signaling; d) Notch signaling; e) Longevity.

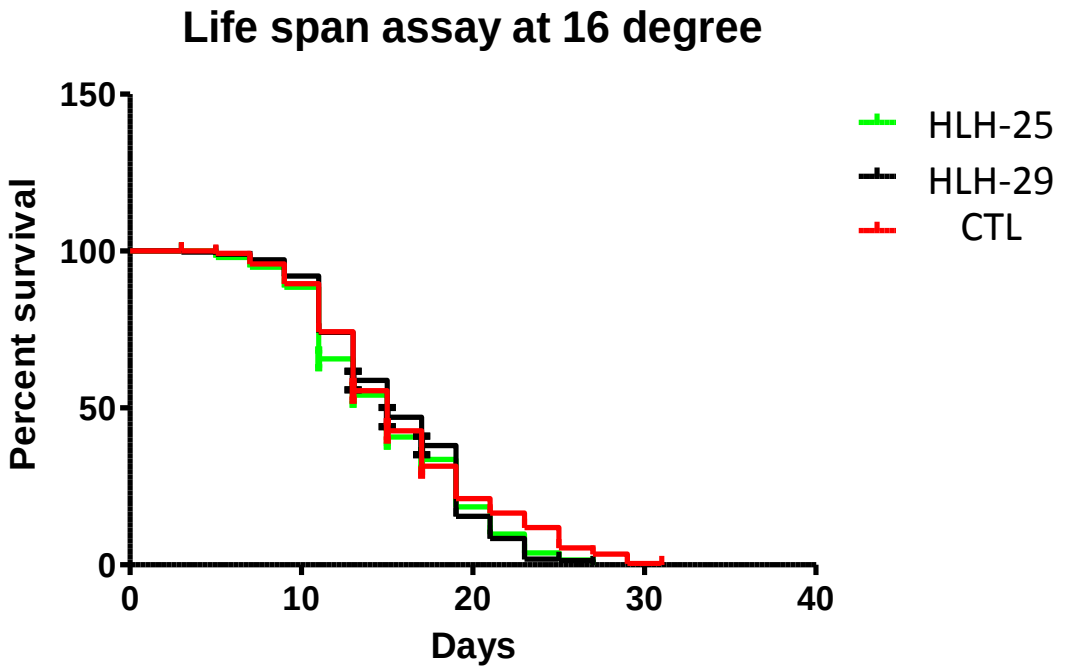

Figure 3.4 Life-span assay. Wild-type animals were growth on NGM plated contained bacteria that expressed wild-type, *hlh-25* and *hlh-29* dsRNA.

4 DISCUSSION

The experiments in this paper lay the foundation for understanding the roles of REF-1 family proteins during *C. elegans* development. Full-genome DNA microarrays were used to profile transcript changes caused by loss of *hlh-25* or *hlh-29*. Our results show that, 81% of 634 *hlh-25*-regulated genes showed increased expression (Table S1) and 88% of 284 regulated genes showed increased expression in *hlh-29* mutant animals (Table S1), which indicate that HLH-25 and HLH-29 mainly act as repressors to regulated transcription and lead a hypothesis that REF-1 family primarily regulates the development as suppressors.

Additionally, functional analysis based on the GO database of differentially expressed genes in *hlh-25*-mutant animals revealed changes in genes connected to embryonic development, nematode larval development, and regulation of growth. Our data supports previous finding that HLH-25 is likely to regulate developments in embryos and larval (Grove et al., 2009). In the previous study, *hlh-29* was predicted to affect genes involved in signaling, locomotion, reproduction and translation. In our study, *hlh-29*-regulated genes, unexpectedly, were not connected to these GO terms, but interestingly, contributed to "positive regulation of growth" and "aging" (Table 3.3) suggesting that HLH-29 may have a role in regulating growth and longevity.

Network studies provide further understanding of molecular functions of proteins. Our network analysis of HLH-25 led to the identification of several networks that have overlapping functions in cell divisions. This result strongly supports the previous study which shows that the "cell division" GO term is only associated with two bHLH proteins, MDL/MXL and HLH-25, among the whole bHLH protein family (Grove et al., 2009). Additionally, identifying these networks allowed us to see relationships between genes involved in cell division and Notch signaling and to show how REF-1 family proteins and HES proteins in general can act to transduce

these signals. Further work in the lab could focus on how REF-1 family proteins affect the cell division process.

Our data also provides the first evidence that the REF-1 proteins may function in Wnt signaling. The Wnt signaling pathway plays a pivotal role in regulating development in many diverse organisms, including *Drosophila, C. elegans* and humans (Wodarz and Nusse, 1998; Hobmayer et al., 2000; Peifer and Polakis, 2000). Developmental events such as cell proliferation, polarity, fate specification, migration of cells and differentiation are regulated by this pathway (Hobmayer and et al., 2000), and this regulation can occur via the canonical Wnt/ β-catenin pathway or the β-catenin independent (non-canonical) pathways (Komiya, 2008). In *C. elegans*, the β-catenin independent pathways mediate the signal from P2 cell to the EMS cell and control the processes of T cell polarity and Z1/Z4 polarity (Rocheleau et al., 1999; Smit et al., 2004; Herman, 2001; Siegfried et al., 2004). DSH-2 and LIT-1 are two components involved in β-catenin independent pathways and function to regulate cell fate specification and direct embryonic polarity, respectively. Future studies in the lab can exploit the relationship between HLH-25 and non-canonical Wnt signaling to better understand how DSH-1 and LIT-1 function.

Notch signaling pathway plays a prominent role in mediating cell-cell interactions during animal development. In *C.elegans*, GLP-1 is one of the two Notch trans-membrane receptors (Farley et al.2010). By negatively regulating the translational suppressor of GLP-1, POS-1, both HLH-25 and HLH-29 might participate in feedback regulation of Notch signaling. It will be interesting to see if the other REF-1 family proteins also do this.

Multiple pathways are found to affect longevity in *C. elegans* (Kenyon, 2005). For example, insulin/insulin-like growth factor (IGF-1) signaling (IIS) pathways shorten life-span by activating mitochondrial activity (Feng et al, 2001). In the IIS pathways, DAF-2, the insulin/IGF-1 re-

ceptor, is the only receptor of the pathway, and negatively regulates the FOXO transcription factor DAF-16 through a conserved PI-3 kinase/PDK/AKT, SGK cascade (Kenyon, 2005). In the pathway, AKT-2, a serine/thereonine kinase, acts to repress the expression of DAF-16 (Quevedo et al., 2007). In our study, AKT-2 is positively regulated by HLH-25, and HLH-29 shares at least one transcriptional target, SOD-1, with DAF-2 and DAF-16 (Vanfleteren et al. 1995). SOD-1 is a copper/zinc superoxide dismutase which functions to control reactive oxygen species within the cell that can significantly influence life span when its gene expression is altered. Taken together, the REF-1 family may regulate components or targets of the IIS pathways, as well as genes that regulate organismal responses to reactive oxygen species, to affect the longevity in *C. elegans*. Our inability to correlate this molecular data with data from our behavior assays maybe because of the difficult nature of the assay. Future studies in the lab will address this possibility. Alternatively, HLH-25 and HLH-29 may function in only one of several redundant pathways to affect lifespan.

5 CONCLUSIONS

From our results, we conclude that HLH-25 and HLH-29 mainly act as repressors to regulate transcription, and we speculate that this may be a characteristic of the REF-1 family as a whole. We also conclude that HLH-25 may play a major role in linking multiple signals to the overall growth and development of *C. elegans*, such as linking cellular division with Wnt signaling. Finally, we conclude that HLH-25 and HLH-29 may both act in a feedback loop to deregulate Notch signaling via negatively regulating POS-1.

6 REFERENCES

Alifragis, P., Poortinga, G., et al. (1997). "A network of interacting transcriptional regulators involved in Drosophila neural fate specification revealed by the yeast two-hybrid system." Proc. Natl Acad. Sci. 94: 13099–13104.

Alper, S., Kenyon, C. (2001). " REF-1, a protein with two bHLH domains, alters the pattern of cell fusion in C. elegans by regulating Hox protein activity." Development 128: 1793-1804.

Bellanger J.M., Gönczy P.(2003). "TAC-1 and ZYG-9 form a complex that promotes microtubule assembly in C. elegans embryos." Curr Biol.13(17):1488-98.

Cristina, D., Cary, M., Lunceford, A., et al. (2009). "A regulated response to impaired respiration slows behavioral rates and increases lifespan in Caenorhabditis elegans." PLoS Genet 5(4): e1000450.

Dawson, S.R., Turner, D.L., et al. (1995). "Specificity for the Hairy/Enhancer of split Basic helix-loop-helix (bHLH) proteins maps outside the bHLH domain and suggests two separable modes of transcriptional repression." Mol. Cell. Biol. 15(12): 6923–6931.

Del, M.G., Grego-Bessa, J., et al. (2007). "Monitoring Notch1 activity in development: evidence for a feedback regulatory loop." Dev. Dyn. 236(9): 2594-614.

DeRenzo C., Reese K.J., et al.(2003). "Exclusion of germ plasm proteins from somatic lineages by cullin-dependent degradation." Nature 424(6949):685-9.

Farley, Brian, et al. (2010). "Integration of glp-1 post-transcriptional regulation in the germline." Develop and Gene Expression

Feng W. M., Sun W. G. and et al. (2001). "Skeletalization characters, classification and evolutionary significance of Early Cambrian Monoplacophoran maikhanellids." Acta Palaeontologica Sinica 40:195-213.

Gaiano, N., Fishell,G. (2002). "The role of notch in promoting glial and neural stem cell fates." Ann.Rev. Neurosci. 25: 471–90.

Geles K.G., Johnson J.J., et al.(2002). "A role for Caenorhabditis elegans importin IMA-2 in germ line and embryonic mitosis." Mol Biol Cell 13:3138-47.

Giagtzoglou, N., Alifragis, P., et al. (2003). "Two modes of recruitment of E(spl) repressors onto target genes." Development 130: 259-270.

Good, K.mR. Ciosk, et al. (2004). The T-box transcriptional factors TBX-37 and TBX-38 link GLP-1/Notch signaling to mesoderm induction in C.elegans embryos. Mol Cell Biol 23:6922-6935.

Grove, C. A., Masi, F. D., et al. (2009). "A Multiparameter Network Reveals Extensive Divergence between C. elegans bHLH Transcription Factors." Cell 138(2): 314-327.

Henderson S.T., Bonafè M., et al.(2006). "daf-16 protects the nematode Caenorhabditis elegans during food deprivation." J Gerontol A Biol Sci Med Sci 61(5):444-60.

Herman M. **(2001). "**C. elegans POP-1/TCF functions in a canonical Wnt pathway that controls cell migration and in a noncanonical Wnt pathway that controls cell polarity." Development 128:581–590.

Hobmayer, B., Rentzsch, F. and et al. (2000). "Wnt signalling molecules act in axis formation in the diploblastic metazoan hydra." Nature 407: 186-189.

Jennings, B., Preiss, A., et al. (1994). "The Notch signalling pathway is required for Enhancer of split bHLH protein expression during neurogenesis in the Drosophila embryo." Development 120 (12):3537-3548.

Joshua, M., Kerry, B., et al. (2003). "Transcriptional outputs of the Caenorhabditis elegans forkhead protein DAF-16." Aging Cell 2:111-121.

Kamath, R.S., Ahringer, J. (2003). "Genome-wide RNAi screening in Caenorhabditis elegans." Methods 30:313–321.

Kashyap, Luv, et al.(2009). " Role of daf-2/insulin signaling pathway genes on preservation of mobility and muscle function in Caenorhabditis elegans ." International Worm Meeting .

Kenyon, C. (2005). "The plasticity of aging: insights from long-lived mutants." Cell 120: 449–460.

King, I.N., Kathiriya, I. S. et al.(2006). "Hrt and Hes negatively regulate Notch signaling through

interactions with RBP-Jj." Biochemical and Biophysical Reasearch Communication. 345:446-452.

Komiya, Y., and Raymond A. (2008) "Wnt signal transduction pathways." Organogenesis 4.2 : 68-75.

Kostrouchova, M., Krause, M., et al. (2001). "Nuclear hormone receptor CHR3 is a critical regulator of all four larval molts of the nematode Caenorhabditis elegans." Proc. Natl. Acad. Sci. 98(13): 7360–7365.

Lanjuin, A., Claggett, J., et al. (2005). "Regulation of neuronal lineage decisions by the HES-related bHLH protein REF-1." Dev Biol. 290(1): 139-51.

Larsen, P. L., Clarke, C. F. (2002). "Extension of life-span in Caenorhabditis elegans by a diet lacking coenzyme Q." Science 295: 120–123.

Leimeister, C., Dale, K., et al. (2000). "Oscillating expression of c-Hey2 in the presomitic mesoderm suggests that the segmentation clock may use combinatorial signaling through multiple interacting bHLH factors." Dev Biol. 227: 91–103.

Liu, Z. J., Shirakawa, T., et al. (2003). "Regulation of Notch1 and Dll4 by vascular endothelial growth factor in arterial endothelial cells: implications for modulating arteriogenesis and angiogenesis." Mol. Cell. Biol. 23(1): 14-25.

Massari, M. E., Murre, C. (2000). "Helix-Loop-Helix Proteins: Regulators of Transcription in Eukaryotic Organisms." Mol. Cell. Biol. 12: 429–440.

McMiller, T. C., Sims, D., et al. (2007). "Molecular characterization of the Caenorhabditis elegans REF-1 family member, hlh-29/hlh-28." Biochim. Biophys. Acta. 1769: 5-19.

Neves, A., Priess, J. R. (2005). "The REF-1 Family of bHLH Transcription Factors Pattern C. elegans Embryos through Notch-Dependent and Notch-Independent Pathways." Dev. Cell 8: 867–879.

Ninkovic, J., Tallafuss, A., et al. (2003). "Inhibition of neurogenesis at the zebrafish midbrain-hindbrain boundary by the combined and dose-dependent activity of a new hairy/E(spl) gene pair." Development 132:75-88.

Peifer, M., Polakis, P. (2000). "Wnt signaling in oncogenesis and embryogenesis--a look outside the nucleus." Science 287(5458): 1606--1609.

Pierce, S.B., Costa, M., et al. (2001). "Regulation of DAF-2 receptor signaling by human insulin and ins-1, a member of the unusually large and diverse C.elegans insulin gene family." Genes & Dev 15: 672–686.

Polinko E., Korf I., et al.(1997)."The Role of CKS-1 in Embryogenesis." International C. elegans Meeting .

Priess,J.R. (2005). "Notch signaling in the C. elegans embryo, In C. e. R. Community (Ed.)." Wormbook 10.

Quevedo C., Kaplan D.R., et al.(2007). "AKT-1 regulates DNA-damage-induced germline apoptosis in C. elegans." Curr Biol.17(3):286-92.

Rocheleau C.E., Yasuda J. and et al. (1999) "WRM-1 activates the LIT-1 protein kinase to transduce anterior/posterior polarity signals in C. elegans." Cell. 97:717–726.

Schubert C.M., Lin R., et al.(2000). "MEX-5 and MEX-6 function to establish soma/germline asymmetry in early C. elegans embryos." Mol Cell 5:671-82.

Severance, S., Rajagopal, A., et al. (2010). "Genome-wide analysis reveals novel genes essential for heme homeostasis in Caenorhabditis elegans." PLoS Genet 6(7):e1001044.

Shelton C.A., Bowerman B. (1996). "Time-dependent responses to glp-1-mediated inductions in early C. elegans embryos." Development. 122(7):2043-50.

Siegfried K.R., Kidd A.R. and et al. (2004) **"**The *sys-1* and *sys-3* genes cooperate with Wnt signaling to establish the proximal-distal axis of the *Caenorhabditis elegans* gonad." Genetics. 166:171–186.

Smit L., Baas A. and et al. (2004) "Wnt activates the Tak1/Nemo-like kinase pathway." J. Biol. Chem. 279:17232–17240.

Sölter, M., Locker, M., et al. (2006). "Characterization and function of the bHLH-O protein XHes2: insight into the mechanisms controlling retinal cell fate decision." Development 133: 4097-4108.

Srayko M., Kaya A., et al.(2005). "Identification and characterization of factors required for microtubule growth and nucleation in the early C. elegans embryo." Dev Cell 9:223-36.

Sulston, J., Hodgkin,J. (1988). "In Wood,W.B. (ed.), The Nematode Caenorhabditis elegans." Methods 587-606.

Tabara H., Hill R.J., et al.(1999). "pos-1 encodes a cytoplasmic zinc-finger protein essential for germline specification in C. elegans." Development 126:1-11.

Walhout A.J., Reboul J., et al.(2002)."Integrating interactome, phenome, and transcriptome mapping data for the C. elegans germline." Curr Biol. 12(22):1952-8.

Wodarz A., Nusse R. (1998). "Mechanisms of Wnt signaling in development." Annu. Rev. Cell Dev. Biol. 14:59–88.

Vanfleteren J.R., Vreese A.D.(1995). "age-1 AND daf-2 MODULATE METABOLIC ACTIV-ITY." International C. elegans Meeting