ABSTRACT

JOHN MARSHALL
Microarray Analysis of the Role of O-GlcNAc in DAF-2 Signaling in *Caenorhabditis elegans*
(Under the Direction of DR. LANCE WELLS)

Mutations in genes functioning along the insulin-like signaling pathway in *Caenorhabditis elegans* extend life span up to three times the wild type phenotype. We sought to explain the mechanism behind this phenomenon by documenting gene expression in populations with these mutations via microarray analysis. The DAF-2 protein is analogous to the mammalian insulin receptor and functions at the head of the complex insulin-like signaling pathway. One branch of the pathway functions to antagonize DAF-16, a transcription factor which mediates hypomorph *daf-2* phenotypes. These phenotypes include extended median lifespan, resistance to oxidative stress, increased longevity, thermotolerance, reduced fecundity, and delayed post-embryonic development. The first two of these are reported as O-GlcNAc dependent. O-GlcNAc is a common regulatory post-translational modification to proteins along this pathway. Microarrays were performed on six different mutant populations. Because in the *daf-16; daf-2* mutants all six phenotypes are brought back to wild type and in the *ogt-1; daf-2* mutant only the first two are, we were able to generate lists of candidate genes impacting each subset of phenotypes by microarray comparisons. This study is important to the genetic investigation of lifespan and carries implications for any species with the insulin signaling pathway.

INDEX WORDS: DAF-2, DAF-16, O-GlcNAc, Life span, Microarray, Insulin-like signaling

MICROARRAY ANALYSIS OF THE ROLE OF O-GlcNAc IN DAF-2 SIGNALING IN *Caenorhabditis elegans*

by

JOHN BENJAMIN MARSHALL

A Thesis Submitted to the Honors Council of the University of Georgia
in Partial Fulfillment of the Requirements for the Degree

BACHELOR OF SCIENCE

in BIOCHEMISTRY AND MOLECULAR BIOLOGY

with HIGHEST HONORS

and CURO SCHOLAR DISTINCTION

Athens, Georgia

2009

MICROARRAY ANALYSIS OF THE ROLE OF O-GlcNAc IN DAF-2 SIGNALING IN

Caenorhabditis elegans

by

JOHN BENJAMIN MARSHALL

Approved:

Dr. Lance Wells — 12/18/09
Dr. Lance Wells — Date
Faculty Research Mentor

Approved:

Dr. Michael Tiemeyer — 12/18/09
Dr. Michael Tiemeyer — Date
Reader

Approved:

Dr. David S. Williams — 12/16/09
Dr. David S. Williams — Date
Director, Honors Program, Foundation Fellows and
Center for Undergraduate Research Opportunities

Approved:

Dr. Pamela B. Kleiber — 12/16/09
Dr. Pamela B. Kleiber — Date
Associate Director, Honors Program and
Center for Undergraduate Research Opportunities

ACKNOWLEDGEMENTS

I thank Dr. Lance Wells for mentoring me, teaching me the principles behind this project, and trusting me with such responsibility in his lab. I thank Mrs. Enas El Karim for her patient guidance when teaching me the foundations of the above protocols and for giving me a sense of how to be a part of such a lab.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER 1
INTRODUCTION

As a traditional model organism, *Caenorhabditis elegans* is a well-studied, free-living nematode which feeds on bacteria. In the laboratory, *C. elegans* are grown on nematode growth medium seeded with OP50 *E. coli*. Its genome of approximately 5000 genes is about half the size of that of *Drosophila melanogaster*, and it has long been sequenced. Its life cycle is simple yet precise in timing when kept at stable conditions. At 25˚C, embryogenesis is complete 14 hours after fertilization (7 hours of cell proliferation and organogenesis, and 7 hours more of morphogenesis). Three hours after fertilization, eggs are deposited from the adult, and L1 larvae hatch from the eggs after completion of embryogenesis. Larvae then molt through four larval stages before becoming an adult. The molt to L2 comes 12 hours after hatching, followed by L3 7 hours later and L4 7 hours after that. If starved, worms can molt to a long-lived, sexually inactive, dormant dauer form from L2, and improving environmental stimuli will cause the dauer form to molt directly to L4. Once at the L4 stage, larvae will molt to an adult in about 9.5 hours. By the time it reaches the L4/adult transition, a hermaphrodite has completed both gonadogenesis and sperm production (males will not be addressed as environmental conditions in this study will prevent their development). As an adult, egg production begins, and the opening of the vulva allows selfing. An adult remains fertile for 4 days, and it can survive 10 to 15 days after this. As a worm ages, it displays decreased mobility, alterations in morphology and enzyme activity, and accumulation of the pigment lipofuscin. A normal lifespan is around 20 days (Wood & Community, 1988).

A worm with a *daf-2* loss-of-function mutation has a lifespan up to threefold that of wild type (Lee et al., 2003). A mutation in DAF-16 completely suppresses the extended lifespan of the *daf-2* mutant (Figure 1) (Kenyon et al., 1993). These mutations are found in the insulin-like pathway in *C. elegans*, termed the dauer formation pathway (*daf*) because it was first discovered as mediating the molt to the long-lived, sexually immature dauer form in response to nutrient availability. Kenyon et al. (1993) first proposed that the longevity of the dauer is not due to the state of dormancy but to an unknown mechanism uncoupled from the dauer form in a *daf-2* mutant. Lee et al. (2003) later elucidated this mechanism: DAF-16 encodes a transcription factor which operates only under depressed DAF-2 activity and can up- or down-regulate a variety of gene targets. In other words, DAF-2 signaling must antagonize DAF-16. This provided a convenient explanation for the observed phenomenon. In the *daf-2* mutant, the DAF-16 transcription factor is not antagonized, and the population exhibits extended lifespan. However, in *daf-2; daf-16* animals, DAF-16 is not functional and cannot mediate a response, just as in wild type. Therefore, the double mutant shares the lifespan statistics of wild type.

In general, the DAF-2 protein is analogous to the mammalian insulin receptor, and it functions at the head of a pathway analogous to mammalian insulin signaling (Figure 2). This pathway is complex in any organism and mediates a number of responses. In *C. elegans* DAF-2 signaling mediates dauer formation, processes of nutrient homeostasis (including GLUT-4 vesicle translocation for glucose uptake, glycogen synthesis, and other processes hallmark to the functions of insulin), and it has been implicated in reproductive development, stress response, and immunity (Lizcano & Alessi, 2002; Murphy et al., 2003).

The pathway from DAF-2 to DAF-16 is now known in more detail (Figure 3). DAF-2 is a receptor tyrosine kinase which dimerizes and autophosphorylates upon ligand binding. IST-1

(analogous to the mammalian insulin receptor substrates, IRSs), which contains the phosphotyrosine-binding (PTB) domain, binds to specific phosphotyrosine residues on the receptor, and the receptor in turn phosphorylates IST-1. This allows association of IST-1 with the adaptor subunit (p85) of phosphoinositide 3-OH kinase (PI3K, called AGE-1 in *C. elegans*). The catalytic subunit of PI3K (p110) then phosphorylates the membrane phospholipid phosphatidylinositol (4,5)-bisphosphate to phosphatidylinositol (3,4,5)-triphosphate (PIP3). PIP3 recruits several proteins with pleckstrin homology domains to the membrane, including AKT (also called protein kinase B, PKB) and 3-phosphoinositide-dependent protein kinase-1 (PDK-1). This co-localization of enzyme and substrate allows PDK-1, along with several other proteins functioning as PDK-2, to phosphorylate two different residues on AKT (Lizcano & Alessi, 2002). This activates AKT to then mediate a number of responses specific to insulin, including the phosphorylation and sequestration of DAF-16 in the cytosol (Cahill et al., 2001). DAF-16 is a FoxO1-like transcription factor which is essential for the division of insulin-signaling responses particularly significant to life span, including reproductive development, energy storage, stress response, and immunity (Murphy et al., 2003). Via another unknown mechanism, PI3K (AGE-1) also plays a role in activating serum glucocorticoid kinase (SGK), which functionally converges with the AKT pathway to antagonize DAF-16 (Lizcano & Alessi, 2002).

Therefore, insulin-like signaling suppresses DAF-16 activity. The more general conclusion about lifespan is that hypomorph mutations along the pathway increase lifespan by allowing DAF-16 localization in the nucleus to impact transcription of several gene targets. This includes hypomorphs of DAF-2, AGE-1, and SGK-1, and the extended lifespan in each can be brought back to wild type by adding a hypomorph of DAF-16 (Hertweck et al., 2004; Ogg et al., 1997). Indeed, hypomorphs of both DAF-16 and another gene, PTEN, decrease lifespan (Figure

2 for summary). PTEN (DAF-18 in *C. elegans*) functions along this pathway as a phosphatase, counter to the activity PI3K (AGE-1). It cleaves a phosphate from PIP3, thereby reducing the concentration of the membrane anchor of AKT and PDK-1 (Lizcano & Alessi, 2002). Therefore, a hypomorph in DAF-18 cannot counterbalance PI3K, and the flux through the pathway is increased, preventing DAF-16 nuclear localization, thereby reducing lifespan.

In unpublished data, Wells (2008) has demonstrated the importance of *N*-acetylglucosamine (O-GlcNAc) in modulating lifespan via fine-tuning of the insulin signaling pathway. O-GlcNAc is a common post-translational modification which is dynamic, ubiquitous, and essential (Wells et al., 2001). A connection between O-GlcNAc and insulin signaling has been implied by investigations of insulin resistance. Though the molecular mechanism remains obscure, flux through the hexosamine biosynthetic pathway has been strongly implicated in the reduction in potency of insulin signaling (Wells et al., 2003). Part of the total glucose uptake (mediated by insulin signaling) is diverted through the hexosamine biosynthetic pathway (HSP) to produce uridine 5'-diphospho *N*-acetylglucosamine (UDP-GlcNAc). UDP-GlcNAc is highly sensitive to ambient glucose levels. Because UDP-GlcNAc is also the donor sugar nucleotide for the O-GlcNAc modification, a model has been proposed of O-GlcNAc as a nutrient sensor (Wells et al., 2001). Increased flux through the HSP leads to elevated O-GlcNAc modification to a number of proteins, now known to include those along the insulin pathway (Wells unpublished data). Elevated global O-GlcNAc in turn causes insulin resistance by inhibiting glucose uptake via a mechanism not yet elucidated (Wells et al., 2003).

The two O-GlcNAc cycling enzymes, O-GlcNAc transferase (OGT, which adds the modification) and O-GlcNAcase (OGA, which cleaves the modification) are both cytosolic and nuclear enzymes (Iyer & Hart, 2003). Most species contain only one gene for each. Increasing

evidence is for a ying-yang interplay between phosphorylation and O-GlcNAcylation in regulatory processes (Hart, 1997). In humans, OGT maps to Xq13 and OGA maps to 10q24, both of which are associated with age-related neurological disorders (Shafi et al., 2000; Gao et al., 2001). The *C. elegans* genome contains one ortholog for OGT, OGT-1, and one for OGA, OGA-1.

The connection between O-GlcNAc and insulin signaling has been solidified by considering the lifespan phenotype. Wells (unpublished data) has indicated that OGA and OGT modulate lifespan. An *oga-1* mutant (higher amounts of O-GlcNAc-modified proteins) appears wild type except for an elevated median lifespan for the population (approximately 75 percent of the worms are alive at the median wild-type age). An *ogt-1* population (lower amounts of O-GlcNAc-modified proteins) is likewise fertile but demonstrates reduced median lifespan, with approximately 25 percent alive at the wild-type median. These correlations were confirmed as causative by consideration of the *oga-1; ogt-1* double mutant population, which has a lifespan more similar to the *ogt-1* population than *oga-1* (Figure 4A).

These lifespan modulations were found to be DAF-16-dependent (Figure 4B), indicating that the mechanism for modulation was via the insulin-like-signaling pathway upstream of DAF-16. Further double mutant studies pinpointed the location of action along this pathway. Worms with a *daf-18* null allele have reduced lifespan due to activation of the signaling cascade and inhibition of DAF-16. The *daf-18; oga-1* double mutant lowers the *oga-1* extended lifespan phenotype, indicating that O-GlcNAc is acting upstream of DAF-18. Also, addition of an *ogt-1* allele to *age-1* or *sgk-1* mutants suppresses the extended lifespan in either case, indicating that O-GlcNAc must be also functioning downstream of these effector kinases. These studies limit O-

GlcNAc modification to insulin-like signaling between DAF-2 and PDK-1, as well as along the SGK branch, which diverges at AGE-1 and converges at AKT.

Interestingly, while addition of *daf-16* to *daf-2* brings maximal longevity back to wild type, addition of *ogt-1* to *daf-2* only brings back median lifespan (Figure 4C). This allowed us to use microarray analysis to determine which genes were being brought back to wild-type regulation from *daf-2* levels in the *daf-16; daf-2* double mutant versus those brought back in the *ogt-1; daf-2* mutant. The first set must include genes mediating the return of both median life span and maximal longevity back to wild-type levels, and the second set must include only those mediating the return of median life span. By breaking down the phenotypic gap between the *daf-2* and *daf-16; daf-2* mutants, the *ogt-1* mutation allowed a more specific approach to the investigation of the genetic basis of longevity.

In the broader picture of all *daf-2* phenotypes, lifespan has been correlated to alterations in reproduction and development, as well as oxidative and thermal stress resistance (Murphy et al., 2003). Wells (unpublished data) has shown that some of these phenotypes are O-GlcNAc-dependent and others O-GlcNAc-independent. Of the six *daf-2* phenotypes (extended median lifespan, resistance to oxidative stress, increased longevity, thermotolerance, reduced fecundity, delayed post-embryonic development), the first two are brought back to wild type with the addition of an *oga-1* allele. The purpose of this study is to explain these results.

What exactly is the mechanism by which these mutations are modulating lifespan? Though the role of DAF-16 is now better understood as a transcription factor, what is not understood is the dichotomy between those phenotypes mediated by the insulin-like signaling pathway which can be brought back to wild type by mutations influencing levels of O-GlcNAc modification versus those phenotypes which seem to be O-GlcNAc-independent. To help

elucidate these mechanisms, microarrays were performed on seven *C. elegans* populations, including N2 (wild type) and populations containing the following mutations or combinations of mutations:

RB653, *ogt-1(ok430)*
RB1169, *oga-1(ok1207)*
CF1041, *daf-2(e1370)*
CF1380, *daf-16(mu86)*
daf-16(mu86); daf-2(e1370), and
ogt-1(ok430); daf-2(e1370).

By comparing results from these microarrays, we were able to generate the hypothesized lists of genes potentially mediating the O-GlcNAc-dependent *daf-2* phenotypes, as well as those potentially mediating O-GlcNAc-independent phenotypes. Importantly, lifespan is among those phenotypes, and this study therefore provided a ground work for further study in the field.

Additionally, by virtue of having microarrays of both *ogt-1* and *oga-1*, this project cast light on the general functions of O-GlcNAc. We hypothesized the discovery of genes reciprocally regulated in each case. This hypothesis was not supported, and instead several genes were found to be regulated in a consistent fashion in both mutant populations.

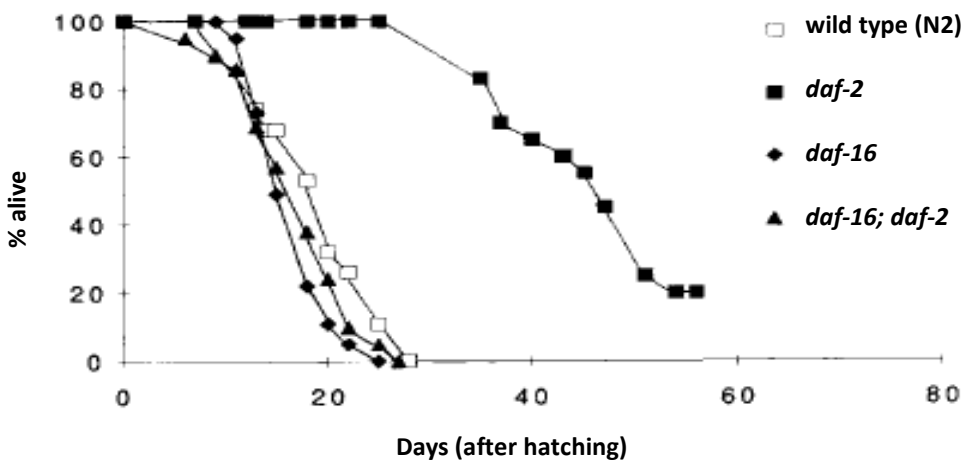

Figure 1: Addition of the *daf-16* mutation to *daf-2* suppresses *daf-2* extended maximal longevity and median lifespan (data adapted from Kenyon et al., 1993).

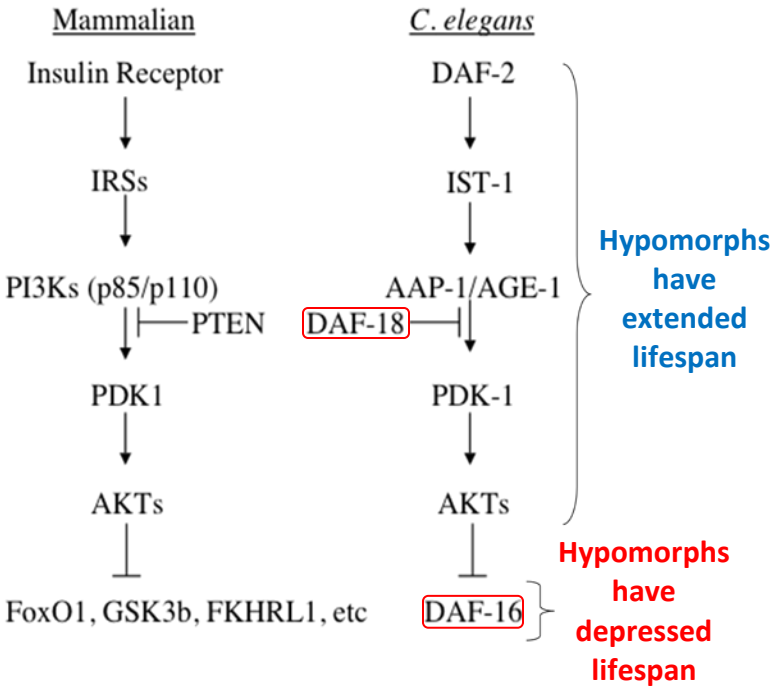

Figure 2: Parallels between the mammalian insulin signal transduction pathway and the analogous pathway in *C. elegans* (Wells unpublished data).

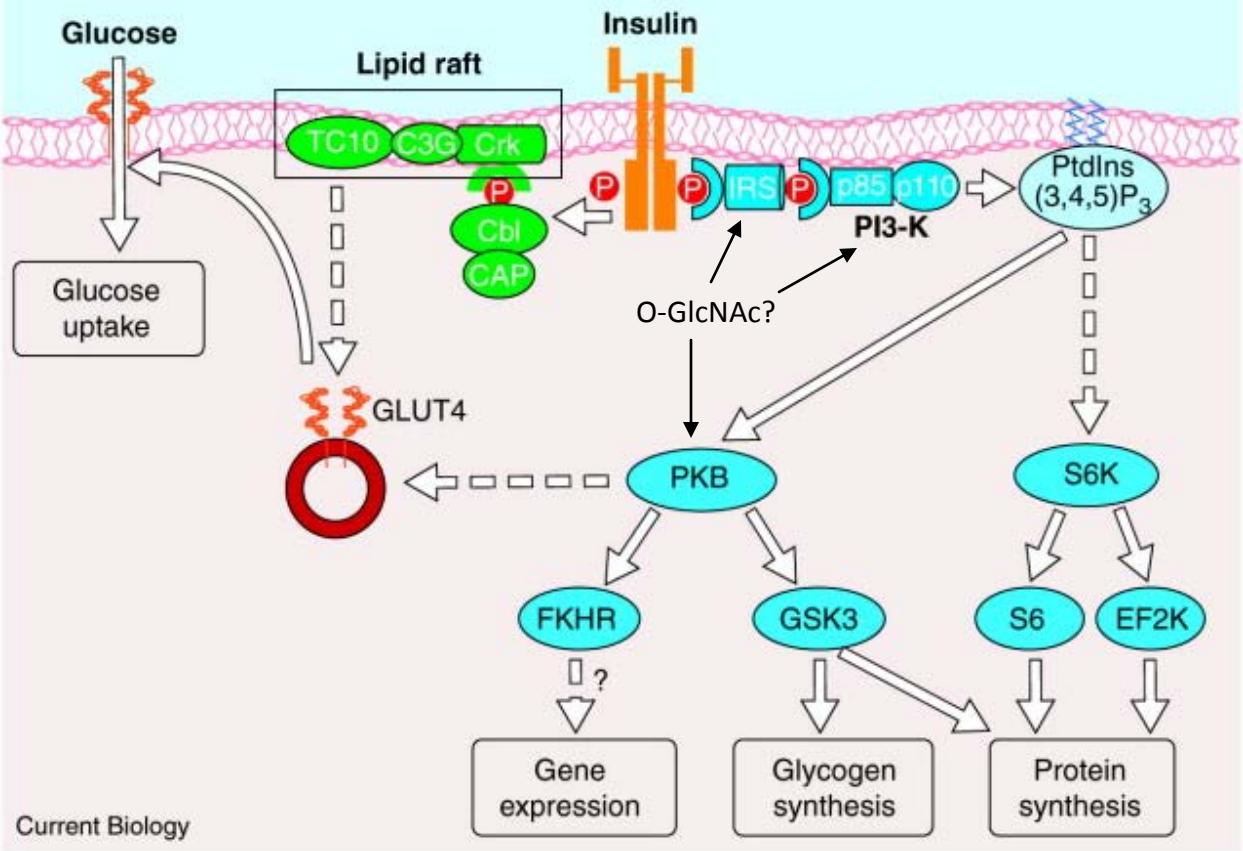

Figure 3: Insulin signaling to PIP3 and recruitment of proteins with pleckstrin homology (not shown, AKT and PDK-1) (modified from Lizcano & Alessi, 2002).

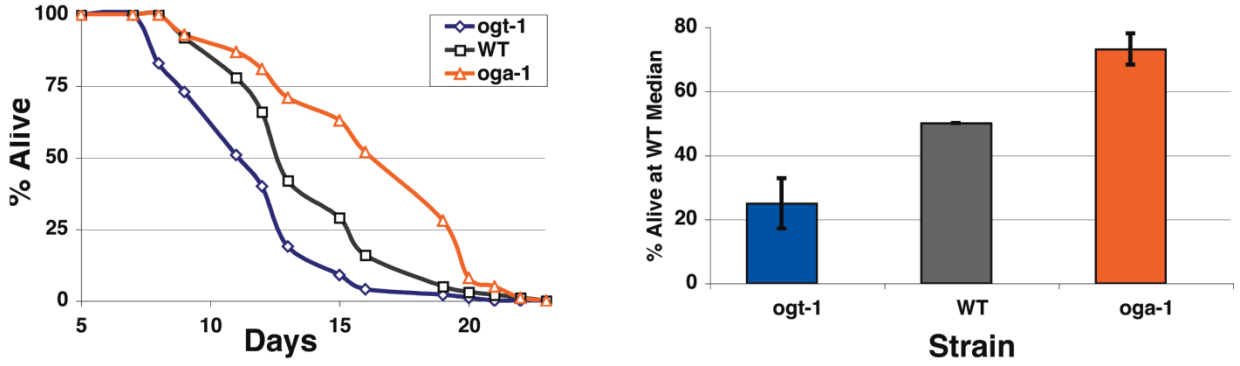

Figure 4A: OGT and OGA modulate life span (Wells unpublished data).

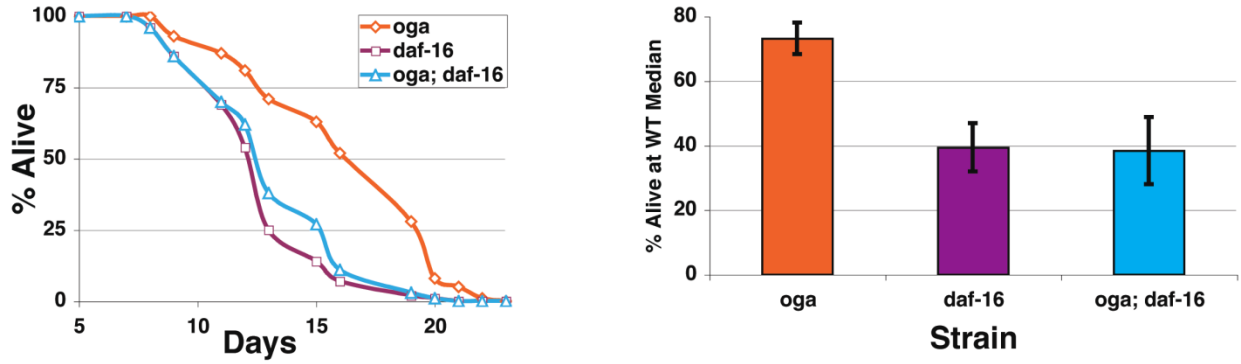

Figure 4B: Lifespan phenotypes of mutations in the O-GlcNAc cycling enzymes are DAF-16-dependent (Wells unpublished data).

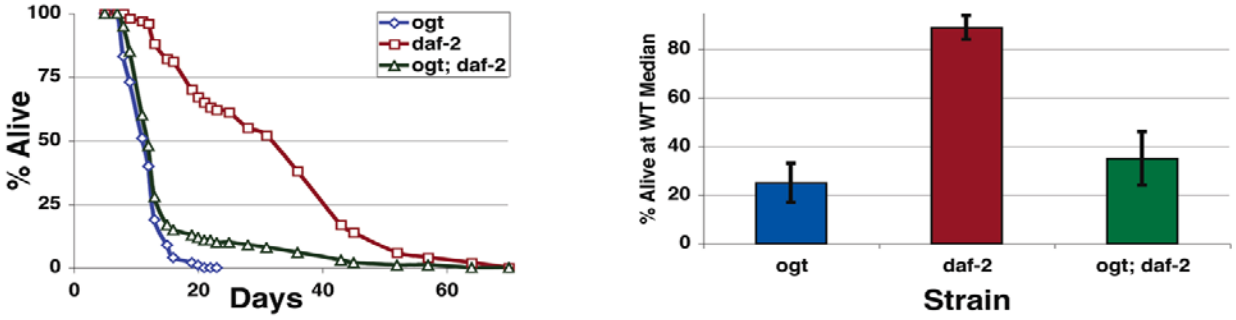

Figure 4C: Adding the *ogt-1* mutation to *daf-2* suppresses median life span but not maximal longevity (Wells unpublished data).

CHAPTER 2
PROCEDURE

Consistent pouring of agar plates was necessary to complete this project. 10 cm plates were needed, which required 3X nematode growth medium (NGM). To make 1 L of NGM, 3 g NaCl, 7.5 g peptone, 20 g agar, 975 ml molecular water, and 1 ml each of cholesterol (5 mg/ml), 0.1 M $CaCl_2$ (stored at 4˚C), and 1 M $MgSO_4$ were combined, dissolved, and autoclaved in a narrow-necked 2 L flask. After cooling for fifteen minutes in a 35˚C water bath, 25 ml KH_2PO_4 (1 M at pH 6) was added while on a gentle stir. 1 L agar would make around 30 plates. These were seeded the following day using a culture of OP50 *E. coli* cells inoculated in LB broth liquid medium. The lawn was grown for a day on the lab bench, and then they were stored in a cold room until use to prevent contamination.

A second type of NGM plate was also needed: those containing the drug 5-fluoro-2'-deoxyuridine (FUDR). This drug is an antineoplastic agent which acts as a DNA-synthesis inhibitor, and FUDR plates were used in this protocol to prevent the hatching of eggs laid by the target population, thereby preventing mixture with progeny. To prepare these plates, an extra step was added before seeding where 0.1 ml 4 mg/ml solution of FUDR in molecular water was spread across the surface of the agar, resulting in a final FUDR concentration of 54 µM (based on 30 ml agar per plate).

The LB broth (both liquid medium and agar) was also prepared on site, according to standard protocol. To ensure isolated OP50 *E. coli* colonies were available for seeding the LB broth liquid medium, standard *E. coli* stock maintenance was observed, including occasional replacement of overgrown plates.

In order for a population of worms to be ready for RNA extraction, each worm had be 9 days old. Sucrose flotation, bleaching, and recovery of the eggs began the synchronous population, and transfer of L4 larvae to FUDR-containing plates prevented the hatching of any eggs laid subsequently (the target population could not mix with its progeny). This protocol was adapted from those described in Portman (2006), in Stiernagle (2006), and in Wolkol et al. (2000), and it was perfected for maximal yield through trial and error.

First, the population had to be allowed to grow to fill at least twenty-five 10-cm plates, keeping in mind the target volume of worms required to isolate RNA (approximately 1 ml). Stock plates were maintained for each strain, and these were used to seed the plates to expand the populations. In accordance with varying life spans among the strains, each strain grew at a different rate. Once the plates were as full as possible with gravid adults, the worms were harvested (at day 0) from the plates into a 15 ml centrifuge tube using M9 solution. After topping off the tube with M9, the collected solution was spun down and washed with M9 until the supernatant was clear and generally free from bacteria. It was then washed again with cold 0.1 M NaCl. After aspirating, the pellet was resuspended to the 5.5 ml level with cold NaCl and allowed to sit on ice for 15 minutes. Throughout this procedure, centrifuging was done at a maximal speed on a clinical centrifuge (4000 rpm) and for times ranging from 1 to 5 minutes based on if worms remained in the supernatant.

This prepared the worms for sucrose flotation. 5.5 ml cold 60 percent sucrose solution was added (or one equivalent of sucrose per equivalent of NaCl-worm solution), bringing the total sucrose concentration to 30 percent. The tube was inverted to mix and then spun at 4000 rpm for 5 min. The result was both a pellet and a light brown layer on top. This sucrose flotation was performed to isolate the gravid adults (found in the surface layer) from the pelleted debris,

juvenile worms, and eggs which had already been laid. Therefore, a flamed, broken Pasteur pipette was used to remove the top brown layer into another centrifuge tube containing 12 ml cold NaCl. The pellet was discarded. Sucrose is harmful to the worms and the eggs they carry, so after washing the top layer four times with cold NaCl to remove residual sucrose, the gravid adults were ready for bleaching.

The purpose of bleaching was to kill the adults and disintegrate their carcasses, leaving eggs which were all the same age. While the bleach had to be strong enough to destroy the carcasses, it also could not be allowed to destroy the eggs inside. Both the intensity of bleach solution and timing of the procedure were used to regulate this. We used a solution with NaOCl 1.8 percent of final volume and NaOH 0.8 N of final volume. 5 ml of this solution was prepared by combining 1.5 ml 6 percent NaOCl, 0.32 ml 12.5 N NaOH and 3.18 ml molecular water, and this was inverted to mix (the volume of alkaline hypochlorite solution used had to be at least 2.5 times the volume of pelleted worms). This was used to resuspend the pellet of washed gravid adults in a new centrifuge tube, which was allowed to incubate on a rocking platform for 5 to 10 minutes, or until the bodies were completely dissolved (closely monitored). As soon as only a cloudy solution of eggs remained, 8.5 ml cold NaCl was added to dilute the bleach and aid in the pelleting of the eggs. The eggs were then spun down for 3.5 minutes and washed four times with M9 to prevent the bleach from harming them. After resuspending the pellet in about 0.5 ml M9, a Pasteur pipette was used to drop the eggs onto one clean, OP50-seeded NGM plate. This concluded day 0.

On day 2 (less than 48 hours later, before the worms reached sexual maturity), the worms were harvested and transferred to seeded NGM plates containing FUDR. For each regular plate

harvested, the population was distributed over at least 15 FUDR plates to ensure that the populations had enough food to last another 7 days.

On day 9, the worms were harvested one final time. M9 was lightly pipetted onto the plate, swirled around, and collected. This is different from the vigorous washing of the plate employed earlier in the protocol, because this method allowed all eggs to stay stuck to the bacteria food. Many worms were lost this way, but this was the most efficient way to surely collect only worms and no eggs. Upon completion of harvesting, the worms were washed until the supernatant was clear and free from bacteria. The product tube containing the aspirated pellet was kept at -80˚C until isolation of the RNA.

The above protocol was repeated for each strain.

As adapted from Portman (2006), RNA isolation included freeze cracking for lysis and DNase treatment to degrade genomic DNA. The protocol began by thawing the samples and adding 5 to 10 volumes of Trizol for each volume of isolated worms. For freeze cracking, the Trizol solution was frozen in liquid nitrogen and then transferred to a 37˚C water bath to completely thaw. This step was repeated several times. To disrupt RNA-protein complexes, a cycle of 30 seconds vigorous vortexing and 30 seconds rest was repeated five times, and the tubes were then allowed to rest for at least 5 minutes. 0.2 ml chloroform was then added for each 1 ml Trizol added, and the tubes were vortexed 30 seconds. They were incubated 2-3 minutes and spun for 15 minutes at 12,000 rpm. This left an upper aqueous layer containing the RNA, which was transferred to a new tube. 0.5 ml isopropanol was added for each 1 ml Trizol added. After mixing and incubating for 10 minutes, the tubes were centrifuged to pellet the RNA. The supernatant was decanted, and the pellet was washed with 75 percent ethanol. The tubes were flicked and shaken to get the gel-like RNA pellet to resuspend. This again was spun down,

decanted, and the pellet was left to air-dry to remove residual ethanol. The pellet was then dissolved in 52 µl RNAse-free water (Tris buffer was also used to help dissolve the pellet). This concluded the freeze-cracking, and the concentration and purity were checked on a spectrophotometer to ensure the RNA would be of the quality required by a microarray.

Until this point, the genomic DNA was mixed with the isolated RNA. Therefore, each sample was subjected to DNase activity. 50 µl of the RNA/DNA samples were combined with 25 µl 10x DNase I buffer, 165 µl molecular water, and 10 µl DNase I, per the Ambion Kit.

The final purification steps were carried out with an RNeasy kit (Qiagen). 250 µl (an equal volume from the DNAse reaction) of 70% ethanol was added to the tubes. The mixture was transferred to a spin column, spun at maximum speed for 15 seconds, and the flow-through was discarded. The RNA was then washed with 700 µl buffer RW1, the flow-through discarded, and then it was washed with 500 µl buffer RPE, and again with 500 µl buffer RPE but this time centrifuging for 2 minutes. After the flow-through was discarded, the column was spun once more to remove residual buffer. The RNA was eluted from the column by transferring it to a fresh 1.5 ml centrifuge tube, adding 50 µl water to the center of the filter, allowing a 1-minute incubation, and spinning the column for 30 seconds at maximum speed. The purity and concentration readings were again taken, this time of the purified RNA product. Throughout this protocol, the water used was RNAse-free.

Worms were saved at two stages (once at the beginning and once at the end of the protocol) for strain confirmation to be carried out after the protocol was completed. Worms were picked into a worm suspension buffer before harvesting to begin the bleaching process and before harvesting the final 9-day-old population. These were kept at -80˚C until use for DNA testing. The worms were thawed, lysed with a lysis buffer, the DNA was isolated, a PCR was run

with the correct primers to confirm each mutation, and a gel was run. The *daf-2* mutation is impossible to identify by this method because it is a single-nucleotide substitution, so the DNA of any strain which included *daf-2* was purified and sent off for sequencing.

CHAPTER 3
RESULTS

Ten microarrays were performed. Two were run for each of the following populations: N2, *ogt-1*, and *oga-1*. Single microarrays were run for each of the following: *daf-2*, *daf-16*, the *daf-16; daf-2* double mutant, and the *ogt-1; daf-2* double mutant. 22,625 gene loci were tested, which is representative of the entire C. elegans genome.

Log base 2 ($\log_2$) analysis was first performed of the fold change of each locus of each data set normalized to wild type, given by

$$\log_2\left(\frac{\textit{quantitative expression of gene in mutant population}}{\textit{quantitative expression of gene in wild type}}\right)$$

The repetitions (*ogt-1* and *oga-1*, both normalized to their respective N2) were then used to establish a value of significance. Overlap was calculated as a percentage for each of a number of values of significance (in increments of $\log_2=0.5$), in search of the value which maximized overlap. For example, setting $\log_2=1$ as significant gave the following four lists of genes: those upregulated by at least a two-fold change from wild type in both repetitions of *ogt-1*, those downregulated by at least a two-fold change in both sets of *ogt-1*, those upregulated in both sets of *oga-1*, and those downregulated in both sets of *oga-1*. Summation of the numbers of genes in each list gives a sense of the number of genes regulated in the same fashion in each repetition. This sum was used as the numerator in a fraction where the denominator was the total number of genes found to be significant (including the number found to overlap in each repetition – the

numerator – plus those found to be regulated differently in each case). In a 100 percent overlap, the numerator would equal the denominator, meaning that no genes were found to be significantly up- or downregulated in one microarray of the *ogt-1* or *oga-1* populations which were not significantly regulated in the same fashion in the other microarray of a population derived from the same lineage. The results indicated that $\log_2=2.5$ should be taken as significant (Table 1). Despite this finding, $\log_2=2$ was used in further analyses. The purpose of this study was to generate a list of genes potentially mediating the observed phenotypic variance among mutants pending further study. Therefore, $\log_2=2$ was considered a sufficient eliminator of noise while maintaining a larger subset of potential genes.

The first analysis was of genes whose regulation was influenced by *ogt-1* or *oga-1* mutations. This was done by comparing the lists of genes irregularly regulated in the *ogt-1* microarray versus *oga-1* using both replicates. Hypothesized was a list of genes inversely regulated by the two hypomorphs consistent in each replicate, meaning these genes could be upregulated by elevated O-GlcNAc protein modification in the cell and downregulated by less O-GlcNAc modification, or vice versa. No such genes were found for the significance level of $\log_2=2$, or even the less stringent $\log_2=1$. Instead, a second data analysis strategy revealed five loci which were all regulated in the same manner in each set with at least a significance level of $\log_2=1$. Three of these were found to be upregulated in both (each being a locus of the same gene), where notably each demonstrated a relatively large fold change, with $\log_2\geq6$ for each locus. The other two loci (of different genes) were downregulated in each case with significance $1\leq\log_2\leq2$ (APPENDIX A).

Next was investigation of the phenotypes of *daf-2*. This analysis distinguished between the hypothesized set of genes mediating the two of the six *daf-2* phenotypes brought back in the

ogt-1; daf-2 double mutant (extended median lifespan and resistance to oxidative stress) from the hypothesized genes mediating the other four of the six which are brought back in the *daf-16; daf-2* double mutant but not in *ogt-1; daf-2* (increased longevity, delayed post-embryonic development, thermotolerance, and reduced fecundity).

The process of elimination began with the broad category of those genes up- or downregulated by $\log_2 \geq 2$ or $\log_2 \leq -2$ in *daf-2*, where the negative indicates downregulation. This gave a list of those genes irregularly regulated by a four-fold change due to abrogation of *C. elegans* insulin-like signaling. Of these, two subsets were considered: those brought back to wild-type regulation or beyond in *ogt-1; daf-2* (at least within $-1 \leq \log_2 \leq 1$ or overshot in the direction opposite from regulation in *daf-2*) and those brought back in *daf-16; daf-2*. Those genes at the intersection of these two subsets provided a list of genes potentially mediating those phenotypes brought back to wild type from *daf-2* in both *daf-16; daf-2* and in *ogt-1; daf-2* (i.e. extended median lifespan and resistance to oxidative stress) (APPENDIX B). Those genes brought back to wild type regulation in *daf-16; daf-2* but not in *ogt-1; daf-2* provided a rather lengthy list of genes potentially mediating the other four phenotypes – those which are O-GlcNAc-independent (i.e. increased longevity, delayed post-embryonic development, thermotolerance, and reduced fecundity) (APPENDIX C).

Table 1: Analysis of consistency of several $\log_2$ values of significance.

Value of Significance	$\frac{\textit{Number of Significant Overlapping Genes}}{\textit{Number of Total Significant Genes}}$	Percentage Overlap
$\log_2$(fold change)=1	166 / 1456	11.40
$\log_2$(fold change)=1.5	70 / 487	14.37
$\log_2$(fold change)=2	45 / 174	25.86
$\log_2$(fold change)=2.5	32 / 94	34.04
$\log_2$(fold change)=3	15 / 63	23.81

CHAPTER 4
DISCUSSION

The purpose of this study was simply to generate lists of candidate genes (those found in APPENDICES A-C). Further study should now be conducted to characterize each gene and narrow the search for genes mediating each individual phenotype via RNA interference testing, quantitative real time PCR, or other means.

Interesting in the list of genes implicated in the phenotypes of *daf-2* is the presence of several gene families, including the C-type lectins, the nuclear hormone receptor family (potentially mediating post-embryonic development and fecundity), and several loci of UDP-glucuronosyl transferase activity. In general, all genes listed have now been implicated not only in dauer formation, but also in the regulation of lifespan and its associated phenotypes. Particularly, those genes of APPENDIX B have now been implicated in the mediation of extended median lifespan and resistance to oxidative stress, and those of APPENDIX C have been implicated in the mediation of increased maximal longevity, delayed post-embryonic development, thermotolerance, and reduced fecundity.

Of note is the finding of five genes regulated significantly in the same fashion in both *oga-1* and *ogt-1* mutants. This may indicate some processes which are essential for both O-GlcNAc addition and cleavage. For O-GlcNAC to fit the model of a nutrient sensor and fine-tuner of insulin signaling, addition and cleavage of O-GlcNAc seemingly might regulate reciprocally; however, perhaps this role would be better fulfilled if abnormal O-GlcNAc levels (either above or below normal physiological level) invoked the same cellular response. Indeed, this data indicates that at least at the level of regulation of gene expression, the functions of

addition and cleavage of O-GlcNAc at various signaling pathways must converge. The genes of our list include a lysozyme (involved in pathogen defense) and a Rho guanyl-nucleotide exchange factor. Further study should explore these findings.

APPENDIX A
GENES REGULATED SIMILARLY IN *ogt-1* AND *oga-1*

Probe ID	*ogt-1* Replicate 1	*ogt-1* Replicate 2	*oga-1* Replicate 1	*oga-1* Replicate 2	Gene Title	Description
3 loci upregulated						
173051_s_at	7.42	8.27	8.38	8.78	F22A3.6a	invertebrate lysozyme
172400_x_at	6.37	7.73	7.24	8.11	F22A3.6a	invertebrate lysozyme
194255_x_at	7.24	7.61	8.11	7.94	F22A3.6a	invertebrate lysozyme
2 loci downregulated						
173906_at	-1.38	-1.39	-1.19	-1.57	Y37A1B.17a	hypothetical protein
187934_at	-1.19	-1.11	-1.31	-1.24	M88.5a	hypothetical protein

APPENDIX B
GENES BROUGHT BACK IN *daf-16; daf-2* AND *ogt-1; daf-2*

175239_at	4.31	0.30	-0.26	F15E11.15a
182498_at	3.90	0.01	0.01	H25K10.1 /// status:Partially_confirmed
180256_at	3.78	0.84	-0.22	SKN-1 Dependent Zygotic transcript /// locus:sdz-24
184116_s_at	3.74	0.85	0.38	T05E12.6
184159_at	3.41	0.36	0.08	T05E12.6
180669_s_at	3.08	0.69	0.73	Y4C6B.6 /// status:Confirmed
184839_at	2.96	0.99	0.74	
191482_at	2.75	-0.23	0.35	F01D4.1 /// locus:ugt-43
190404_s_at	2.49	0.84	0.93	F37B4.7 /// locus:folt-2
183309_at	2.47	0.41	0.10	K03H6.2 /// status:Partially_confirmed
189402_s_at	2.40	-0.28	0.27	Peroxisomal Membrane Protein related
177951_s_at	2.38	0.50	0.17	PDB1.1b
186713_at	2.31	0.29	0.95	ZK512.7 /// status:Confirmed
188660_s_at	2.27	-0.24	0.18	F28A12.4 /// peptidase
179610_s_at	2.25	0.58	0.27	T16G1.5 /// status:Partially_confirmed
175109_at	2.24	0.90	1.00	Y34B4A.9 /// status:Confirmed
194054_s_at	2.10	0.60	0.60	F08A8.4 /// F08A8.3 /// ACYL-COENZYME A OXIDASE, PEROXISOMAL (EC 1.3.3.6) (PALMITOYL-COA OXIDASE) (AOX)
179650_at	2.08	0.55	0.76	K08D8.4b
178918_at	-2.25	-0.95	-0.37	T06D8.1a
173257_s_at	-2.31	-0.89	-0.77	Dauer-upregulated
174748_at	-2.34	-0.94	-0.45	T06D8.1a
183670_at	-2.41	-0.51	-0.45	C18A3.1 /// status:Partially_confirmed

APPENDIX C
GENES BROUGHT BACK IN *daf-16; daf-2* BUT NOT *ogt-1; daf-2*

Probe ID	*daf-2*	*ogt-1; daf-2*	*daf-16; daf-2*	Gene Title
193623_at	5.83	5.75	-0.51	F21F3.3 /// protein-S isoprenylsysteine O-
193262_at	5.03	4.38	-0.17	UNCoordinated /// locus:unc-38
186519_at	4.66	2.01	0.32	F15E11.12 /// status:Confirmed
186801_s_at	4.30	4.22	-0.02	F45D11.16 /// F45D11.14 /// F45D11.15 ///
189469_at	3.76	4.26	-2.39	Cysteine PRotease related /// locus:cpr-2
189419_at	3.73	2.78	-1.15	F15B9.6 /// phospholipase A2
182739_at	3.52	3.70	0.44	F40D4.13 /// status:Partially_confirmed
194159_s_at	3.51	1.63	0.62	Y51F10.7 /// status:Confirmed
173840_s_at	3.47	4.38	0.00	C-type LECtin /// locus:clec-13 /// locus:clec-15
184160_at	3.38	3.33	-0.37	RNase H /// locus:rnh-1.3
173310_at	3.36	1.29	0.31	Y51F10.7 /// status:Confirmed
178927_at	3.19	2.29	0.00	K08D8.5 /// status:Confirmed
188822_at	3.13	1.27	0.16	Acyl CoA DeHydrogenase
192786_s_at	3.13	1.77	0.17	B0286.3 /// saicar synthetase/air carboxlyase
179418_at	3.05	4.28	-1.14	Y6E2A.4 / // status:Partially_confirmed
189646_at	3.03	1.15	0.59	K09C4.5 /// sugar transporter
191066_s_at	3.00	2.36	0.22	UDP-GlucuronosylTransferase
188805_at	2.97	1.68	0.66	C55F2.1b
177920_at	2.92	1.81	0.22	D1086.3 /// status:Confirmed
184103_s_at	2.88	4.47	0.86	F07B7.2 /// W09B7.2 /// status:Predicted
175126_s_at	2.86	1.85	0.11	K08D8.5 /// status:Confirmed
191057_s_at	2.74	1.29	0.44	ZC455.4 /// locus:ugt-6
179452_at	2.71	3.25	-0.61	Y6E2A.4 / // status:Partially_confirmed
194157_s_at	2.71	3.16	-1.15	F53B6.8 /// locus:fipr-26
173109_s_at	2.68	2.71	-0.59	Nuclear Hormone Receptor family /// locus:nhr-146
180376_s_at	2.68	3.12	-0.36	Dauer or Aging adult Overexpression /// locus:dao-2
179858_at	2.66	1.25	0.83	C49C8.5 /// status:Confirmed
192617_at	2.61	1.48	-0.72	SOD (superoxide dismutase) /// locus:sod-3
192407_at	2.56	2.22	-0.38	Glutathione S-Transferase /// locus:gst-4
183027_at	2.55	2.92	-0.18	R07C3.5 /// status:Partially_confirmed
173144_at	2.54	2.06	-0.87	C-type LECtin /// locus:clec-85
190637_s_at	2.53	1.80	0.53	F55B11.1 /// xanthine dehydrogenase

(cont.)

Gene ID	*daf-2*	*ogt-1; daf-2*	*daf-16; daf-2*	Gene Title
174283_s_at	2.50	1.18	0.33	F38B6.4 /// GARS/AIRS/GART
186492_s_at	2.50	1.21	-0.02	F15E11.14 /// F15E11.1 /// status:Partially_confirmed
176668_at	2.48	1.68	0.04	C32B5.1a
181715_s_at	2.47	1.53	0.95	T19B10.2 /// status:Confirmed
182106_at	2.46	2.06	0.79	C34H4.1 /// status:Partially_confirmed
180979_at	2.44	1.33	0.03	GRounDhog (hedgehog-like family) /// locus:grd-5
191951_at	2.44	2.18	-0.50	UDP-GlucuronosylTransferase /// locus:ugt-11
181354_at	2.43	2.37	0.08	F53G2.1 /// status:Partially_confirmed
182081_s_at	2.38	1.48	-0.85	E01A2.8 /// K05F6.11 /// status:Partially_confirmed
191541_at	2.37	1.95	-0.51	CaDmium Responsive /// locus:cdr-2
190344_at	2.36	2.48	-0.28	F26D10.12 /// locus:clec-196
182983_s_at	2.35	1.57	0.12	C32B5.1a /// K02E7.7 /// status:Predicted
176042_s_at	2.32	2.22	-0.14	Nuclear Hormone Receptor family /// locus:nhr-146
187282_at	2.31	2.73	-0.78	F22B7.9 /// status:Confirmed
174347_s_at	2.31	1.13	0.47	C49C8.5 /// status:Confirmed
180616_at	2.30	3.17	-0.56	T22B7.3 /// status:Partially_confirmed
171970_x_at	2.29	3.50	0.33	H39E23.3 /// status:Partially_confirmed
181799_at	2.29	3.85	0.94	F40H7.12 /// status:Partially_confirmed
191863_at	2.26	2.68	0.44	C33F10.7a
177235_s_at	2.24	2.16	-0.54	DH11.2 / // status:Partially_confirmed
178635_at	2.24	1.87	-0.24	PhosphoLipid SCramblas /// locus:scrm-4
180023_at	2.22	2.30	-0.47	T05A7.3 /// status:Partially_confirmed
172241_x_at	2.22	2.61	-0.37	C37A5.8 /// locus:fipr-24
177131_s_at	2.18	1.19	-0.29	Y19D10B.7 /// F15E11.13 /// status:Partially_confirmed /// status:Predicted
190849_at	2.17	1.56	-0.11	UDP-GlucuronosylTransferase /// locus:ugt-41
190114_at	2.16	2.93	-0.04	
176044_at	2.16	1.71	1.00	Y41D4B.16 /// status:Confirmed
185217_at	2.15	1.33	0.11	C32B5.15 /// status:Predicted
174159_at	2.13	2.55	0.54	B0285.7 /// Aminopeptidase
191871_s_at	2.13	1.72	-0.43	CaDmium Responsive /// locus:cdr-4
193024_at	2.12	2.12	-0.37	Nicotinamide Nucleotide Transhydrogenase /// locus:nnt-1
191624_at	2.10	1.85	0.69	
172541_x_at	2.09	2.81	-0.26	MiRP K channel accessory Subunit
193761_at	2.07	1.69	-0.13	Palmitoyl Protein Thioesterase /// locus:ppt-1
180157_at	2.06	2.12	-0.24	F57G12.1 /// status:Confirmed
181552_at	2.05	3.34	-0.20	T22B7.3 /// status:Partially_confirmed

(cont.)

Probe ID	*daf-2*	*ogt-1; daf-2*	*daf-16; daf-2*	Gene Title
184710_at	2.04	1.08	0.08	Y9C9A.16 /// status:Predicted
191421_at	2.04	2.16	-0.26	Seven TM Receptor /// locus:str-31
183207_at	2.03	1.17	0.88	D2096.3 /// status:Confirmed
180212_at	2.02	2.67	-0.80	C44B7.6b
190988_at	2.02	2.22	-1.48	ZC410.5a
188934_at	2.02	1.57	0.48	UDP-GlucuronosylTransferase /// locus:ugt-13
181137_at	2.01	1.00	-0.07	F46G10.1
186273_at	-2.00	-1.79	-0.64	R07C12.1 /// status:Partially_confirmed
187486_at	-2.05	-2.08	-0.35	T BoX family /// locus:tbx-11
177337_at	-2.06	-2.28	0.55	R52.2 / // status:Partially_confirmed
178843_at	-2.07	-5.30	-0.33	F08G5.6 /// status:Confirmed
184121_at	-2.18	-2.03	-0.90	C-type LECtin /// locus:clec-43
193595_at	-2.20	-2.16	-0.54	THaumatiN family /// locus:thn-1
171767_x_at	-2.23	-1.23	0.88	
193588_s_at	-2.25	-2.44	-0.15	THaumatiN family /// locus:thn-2 /// locus:thn-1
186443_at	-2.26	-4.34	1.04	F42A10.7 /// status:Confirmed
179645_s_at	-2.29	-3.42	-0.83	F56A8.8 /// status:Partially_confirmed
172268_x_at	-2.43	-2.99	0.70	Nematode Specific Peptide family, group C /// locus:nspc-16 /// locus:nspc-19 /// locus:nspc-20
179165_at	-2.52	-2.02	0.03	W04A4.2 /// status:Predicted
173281_s_at	-2.61	-1.38	0.74	
181946_at	-2.63	-2.49	0.01	C-type LECtin /// locus:clec-10
190499_s_at	-2.72	-3.65	-0.88	C-type LECtin /// locus:clec-5
186706_at	-2.90	-3.02	0.04	K08D10.10 /// status:Partially_confirmed
186971_at	-2.97	-2.63	0.97	C23G10.11 /// status:Confirmed
181947_s_at	-2.99	-3.29	-0.64	C-type LECtin /// locus:clec-10
193293_at	-3.17	-3.48	0.73	Cysteine PRotease related /// locus:cpr-1
180721_at	-3.20	-2.22	-0.51	C45G7.2 /// locus:ilys-2
192496_at	-3.39	-3.79	-0.45	DeHydrogenases, Short chain /// locus:dhs-26
191515_at	-3.87	-2.98	-0.84	F19C7.2 /// lysosomal carboxypeptidase
183848_at	-3.99	-3.07	-0.25	Downstream Of DAF-16 (regulated by DAF-16) /// locus:dod-23
182838_s_at	-4.00	-4.01	-0.91	T05E12.3 /// status:Confirmed
188538_at	-4.15	-3.35	-0.53	INSulin related /// locus:ins-7
191887_s_at	-4.79	-5.16	1.02	Cysteine PRotease related /// locus:cpr-5

WORKS CITED

Cahil, C. M., Tzivion, G., Nasrin, N., Ogg, S., Dore, J., Ruvkun, G., & Alexander-Bridges, M. (2001). Phosphatidylinositol 3-kinase signaling inhibits DAF-16 DNA binding and function via 14-3-3-dependent and 14-3-3-independent pathways. *J Biol Chem*, 276, 13402-10.

Gao, Y., Wells, L., Comer, F. I., Parker, G. J., & Hart, G. W. (2001). Dynamic O-glycosylation of nuclear and cytosolic proteins: Cloning and characterization of a neutral, cytosolic beta-N-acetylglucosaminidase from human brain. *J Biol Chem* 276, 9838-45.

Hart, G. W. (1997). Dynamic O-linked glycosylation of nuclear and cytoskeletal proteins. *Annu Rev Biochem*, 66, 315-35.

Hertweck, M., Gobel, C., & Baumeister, R. (2004). C. elegans SGK-1 is the critical component in the Akt/PKB kinase complex to control stress response and life span. *Dev Cell*, 6, 577-88.

Iyer, S. P., & Hart, G. W. (2003). Dynamic nuclear and cytoplasmic glycosylation: Enzymes of O-GlcNAc cycling. *Biochemistry*, 42, 2493-99.

Kenyon, C., Chang, J., Gensch, E., Rudner, A., & Tabtlang, R. (1993). A *C. elegans* mutant that lives twice as long as wild type. *Nature*, 366, 461-4.

Lee, S. S., Kennedy, S., Tolonen, A. C., & Ruvkun, G. (2003). Daf-16 target genes that control *C. elegans* life-span and metabolism. *Science*, 300, 644-7.

Lizcano, J. M., & Alessi, D. R. (2002). The insulin signalling pathway. *Curr Biol*, 12(7), R236-8.

Murphy, C. T., McCarroll, S. A., Bargmann, C. I., Fraser, A., Kamath, R. S., Ahringer, J., Li, H., & Kenyon, C. (2003). Genes that act downstream of DAF-16 to influence the lifespan of *Caenorhabditis elegans*. *Nature*, 424, 277-84.

Ogg, S., Paradis, S., Gottlieb, S., Patterson, G. I., Lee, L., Tissenbaum, H. A., & Ruvkun, G. (1997). The fork head transcription factor DAF-16 transduces insulin-like metabolic and longevity signals in C. elegans. *Nature*, 389, 994-9.

Portman, D.S. Profiling *C. elegans* gene expression with DNA microarrays (January 20, 2006), *Wormbook*, ed. The *C. elegans* Research Community, doi/10.1895/ wormbook.1.104.1, http://www.wormbook.org.

Shafi, R., Iyer, S. P., Ellies, L. G., O'Donnell, N., Marek, K. W., Chui, D., Hart, G. W., & Marth, J. D. (2000). The O'GlcNAc transferase gene resides on the X chromosome and is essential for embryonic stem cell viability and mouse ontogeny. *Proc Natl Acad Sci USA*, 97, 5735-9.

Stiernagle, T. Maintenance of *C. elegans* (February 11, 2006), *WormBook*, ed. The *C. elegans* Research Community, doi/10.1895/wormbook.1.101.1, http://www.wormbook.org.

Wood, W. B., & The Community of *C. elegans* Researchers (Eds.). (1988). *The nematode Caenorhabditis elegans*. Cold Spring Harbor Laboratory Press.

Wells, L., Vosseller, K., & Hart, G.W. (2001). Glycosylation of nucleocytoplasmic proteins: Signal transduction and O-GlcNAc. *Science*, 291(5512), 2376-8.

Wells, L., Vosseller, K. & Hart, G.W. (2003). A role for N-acetylglucosamine as a nutrient sensor and mediator of insulin resistance. *Cell Mol Life Sci*, 60, 222-8.

Wolkow, C. A., Kimura, K. D., Lee, M., & Ruvkan, G. (2000). Regulation of *C. elegans* life-span by insulinlike signaling in the nervous system. *Science*, 290(5489), 147-50.