THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

AN EXAMINATION OF LOCOMOTIVE BEHAVIOR, GERMLINE PROLIFERATION, AND LIPID CONTENT IN A NOVEL AUTISM SPECTRUM DISORDER 16P11.2 DELETION *CAENORHABDITIS ELEGANS* MODEL

RHEA ELENA SULLIVAN
SPRING 2018

A thesis
submitted in partial fulfillment
of the requirements
for baccalaureate degree
in Biochemistry and Molecular Biology
with honors in Biochemistry and Molecular Biology

Reviewed and approved* by the following:

Santhosh Girirajan
Assistant Professor of Biochemistry and Molecular Biology
Thesis Supervisor

Lorraine Santy
Associate Professor of Biochemistry and Molecular Biology
Honors Advisor

Wendy Hanna-Rose
Professor of Biochemistry and Molecular Biology
Interim Department Head

*Signatures are on file in the Schreyer Honors College

ABSTRACT

Due to the phenotypic heterogeneity of human neurodevelopmental disorders, it is often difficult to pin point a diagnosis based on behavior alone. Though these disorders have a high degree of comorbidity, there is variance in clinical presentations that can additionally confound diagnosis. Copy number variations (CNV) in the genome have been thought to be the unidentified component that affects such clinical variation. This study examines one rare (<1%) CNV implicated in Autism Spectrum Disorder (ASD), the 16p11.2 recurrent microdeletion. I investigated the individual genes encompassed within this deletion and determined their contributions to the overall phenotypes of motor function, lipogenesis, and embryonic cell line proliferation. *Caenorhabditis elegans* nematode worm is used as a model for its quick reproductive cycle, high-throughput genetic manipulability, and conservation of neural networks.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I met Dr. Santhosh Girirajan on the first day of freshman year through the Schreyer Honors College SHOtime Orientation. Four years later, Dr. Girirajan has mentored me in so many ways that have shaped me to be the person I am today. He has always pushed me to be the best I can be, both in the classroom and at the bench. I am forever grateful to him and Girirajan lab members for their patience and dedication to helping me grow as a student researcher. I'd also like to thank Dr. Wendy Hanna-Rose and her students Dr. Melanie McReynolds and Adam Fenton for being another set mentors to guide me in the world of worms. Additionally, thank you to Dr. Melissa Rolls and students Matthew Shorey and Derek Nye for taking the time out of their busy schedules to teach me fluorescent confocal microscopy.

In the past four years, I was fortunate enough to have not one, but two truly caring advisors. Both Dr. James Endres Howell and Dr. Lorraine Santy have looked out for my best interests, personally and academically.

Most importantly, I'd like to thank Lucilla Pizzo and Ayush Thomas. Lucilla is a big sister both inside and outside the lab. She is the definition of a hard worker and has taught me through her actions that hard work does pays off. And finally, my utmost gratitude goes to Ayush Thomas. Ayush has been by my side since day one and has helped me grow professionally, academically, and personally in the past four years. I don't know where this project would be without him.

Thank you to the Schreyer Honors College, Eberly College of Science, Erickson Discovery Grant, and BMB Department for giving me every opportunity to learn the art of discovery.

CHAPTER 1

INTRODUCTION

Phenotypic Heterogeneity of Neurodevelopmental Disorders

The method in which autism spectrum disorder (ASD) is currently clinically diagnosed represents a need for a paradigm change. Broadly, ASD and other neurodevelopmental disorders, such as schizophrenia (SZ), attention deficit hyperactivity disorder (ADHD), and epilepsy share many of the same behavioral symptoms. This group of conditions shares a common trait of early onset in childhood. Some of these developmental deficits include motor, social, or intellectual impairments.

But yet, even within the same neurodevelopmental diagnosis category as outlined within the Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5)[1], there exists a spectrum of variations and comorbidities. An extreme case of ASD may clinically present as schizophrenia, while a less severe case may present as intellectual disability (ID). Or perhaps, a patient may have ASD comorbid with ID or epilepsy. How must a physician know for sure which symptoms belong to which disorder? A closer look at some of the DSM-5 definitions reveals the overlap of symptoms between disorders. (Table 1).

In reality, no patient exclusively presents all criteria. How do compounded symptoms of epilepsy, severe motor impairments, and macrocephaly/microcephaly confound diagnoses? When coupled with clinical features of ASD, the diagnosis is referred to as "syndromic autism."[2] The phenotypic heterogeneity of neurodevelopmental disorders, like syndromic autism, makes it difficult to create precise symptom categories for each disorder.

Disorder	Criteria A	Criteria B	Criteria C	Specification of Severity
ASD	Abnormal social approach; failure to initiate or respond to social interactions	Restricted, repetitive patterns of behavior, interests or activities	Onset of Criteria A and B during developmental period	Severity is based on social communication impairments and restricted, repetitive behaviors
SZ	Delusions, hallucinations, grossly disorganized behavior	Since onset, level of functioning in various areas of life are severely negatively affected	Continuous signs of disturbance persist for at least 6 months and no attributed to drug abuse	Quantitative assessment of the symptoms of psychosis, including delusions and hallucinations
ID	Deficits in intellectual functions such as problem solving, or planning	Deficits in adaptive functioning that result in failure to meet sociocultural standards	Onset of Criteria A and B during developmental period	Measured by intelligence tests
ADHD	Inattention to details; has difficulty sustaining attention in tasks	Hyperactivity and impulsivity: fidgets or squirms often. Often "on the go."	n/a	Determination if symptoms are in excess and intensity of functional impairment

Table 1. Similarities and Variances Amongst Criteria of Neurodevelopmental Disorders. As defined in the DSM-5, defining characteristics of various neurodevelopmental disorders all included social or intellectual deficits. While presented separately here, often these show comorbidity in patients. Additionally, the severity of such disorders could be affected by other compounded symptoms like macrocephaly, microcephaly, motor defects, and epilepsy.

However, the age of genomics has called for a paradigm shift in how diagnoses of such disorders are categorized. The era of next-generation, high-throughput deep sequencing has changed the pace at which physicians and scientists study such etiologies. The ability to rapidly sequence large amounts of DNA rapidly, cost effectively, and at a greater sequence read depth has led to the discovery of thousands of genetic variants. Computational biology allows for the filtering of vast amount of genomic information. Being able to filter through inheritance models,

types of mutations, presence/absence in control populations, and allelic frequency can help narrow down biomarkers corresponding to each neurodevelopmental disorders[3].

The paradigm of diagnosing neurodevelopmental disorders is shifting from a behavioral assessment to a genome wide screen of variants. The identification of biomarkers allows larger cohorts of children with similar variants, which helps identify their shared clinical features. Additionally, the discovery of these causative variants opens the door for molecular treatments that could prevent certain neurodevelopmental disorders[2].

Effects of Copy Number Variation

Through new developments of chromosomal microarray technologies, hot spots in genome have been identified as conferring significant risk for neurodevelopmental disorders. Some of these hot spots are copy number variances. The identification of copy number variances (CNVs) as potential risk factors have lead scientists to form a rare variant – common disease (RVCD) model of neurodevelopmental disorders[4]. Rare CNVs are stretches of DNA longer than 1Kb that may contain structural variations such as duplications, deletions, and inversions in regions that contain hundreds of genes, disease loci and other functional elements[5]. Though CNVs are rare (<1%) in the population, they account for a significant amount of cases for ID, ASD, SZ epilepsy, and ADHD[6].

Although CNVs have been identified as holding major pathogenic roles, the transition of these identifiers to clinical practice has been slow. There is still no clear molecular mechanism elucidated on how these CNVs interact with each other and with the rest of the genome.

Recurrent 16p11.2 Microdeletion

The 16p11.2 locus has been identified as a major hot spot for neurodevelopmental pathogenicity. Both ~600Kb proximal deletions and duplications have been identified and associated with a range of neurodevelopmental disorders. The deletion is associated with ASD, ID, obesity, and macrocephaly, whereas the duplication is associated with ASD, SZ, and microcephaly[3]. The analysis presented in this paper will focus on the 16p11.2 microdeletion, which represents one of the most common rare CNVs associated with autism to date.

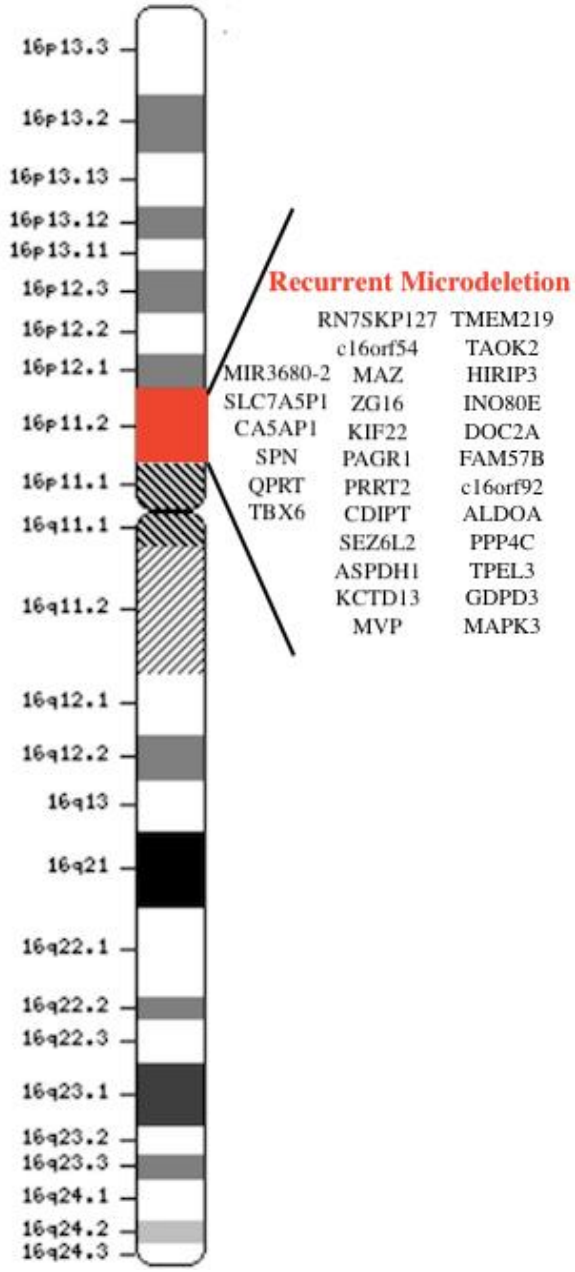

Figure 1. Individual genes encompassed within the deletion. There are over 25 protein-encoding genes deleted in the recurrent microdeletion.

Those with the 16p11.2 deletion may not notice any symptoms, have slight symptoms, or may suffer life altering deficits. Usually it is common for those with the deletion to have developmental delay and intellectual disability. Many also have similar presentations to that of ASD. Other symptoms include impaired communication and socialization. Additionally, some people with this deletion have epilepsy. Physical abnormalities associated with the disorder include are low-set ears and partial syndactyly.

The 16p11.2 deletion is inherited in an autosomal dominant pattern, however; it is most often seen as a *de novo* mutation. Since the CNV is not 100% penetrant, it is hard to predict how rare the deletion is in

the population. Based on GWAS studies, it has been postulated to be 3 in 10,000 people.

There are 29 annotated genes within this region, including transcription factors, chromatin modifiers and genes involved in insulin signaling[3](Fig. 1). However, individual gene contributions to the overall 16p11.2 deletion phenotype have yet to be elucidated. Understanding specific gene contributions and potential gene interactions will help develop potentially more targeted gene therapies.

Untangling Phenotypes from Individual Genes

In order to understand how each individual gene contributes to the overall phenotype of the deletion, each gene will be studied in a *C. elegans* model. For the purpose of this paper, the focus will be on only 7 orthologous genes within this region (Table 2). The first goal is to determine if and how each individual gene (single hit) affects neuromuscular function and proliferation within the mitotic region of the *C. elegans* germline. The second goal is to screen for potential gene interactions by knocking down two genes at a time. This model is meant to emulate the theory behind Dr. Santhosh Girirajan's "Two-Hit" Model. The "Two Hit" model for neurodevelopmental disorders shows that individuals with a second CNV, in addition to the 16p12.1 deletion, will manifest more severe developmental delay.

Though the nematode nervous system is simple, it controls a variety of behaviors, including locomotion[7]. Thrashing, or swimming, has been established as a behavior assay measuring the affect of mutations on movement. Case studies show that probands with the 16p11.2-12.1 deletion suffer delayed motor development and problems with fine motor skills later on in development[8]. Thrashing assays have been used measure neuromotor defects of *C. elegans* models of Alzheimer's Disease[9], Huntington's Disease[10], spinal muscular atrophy,[11] and

Parkinson's Disease[12]. In order to produce high throughput screens, ImageJ wrMTrck plugin was used to quantify the number of body bends in a minute. The body angles through which the worm passes are calculated; the constitution of a bend is when angle of bend and percent pixel area change is above a standard threshold (See Methods).

The second assay performed was staining of *C. elegans* for lipid content. The purpose of staining the animals with Oil-Red O was to rule out an increase in body area from fat content (Fig. 4). *C. elegans* nematodes specifically store fat in the intestinal and skin-like epidermal cells. Key metabolic pathways and their regulators are heavily evolutionarily conserved[13]. Oil-Red O is the best stain to use for staining fat content on *C. elegans* animals, contrary to Nile Red and BODIPY-labeled fatty acids. These stains were originally accepted to dye fat content, until Nile Red was shown to increase in intensity as fasting time increased. Additionally, BODIPY-labeled fatty acids are not *C. elegans* major fat storage compartments. BODIPY has been shown to stain lysosome related organelles, not fat storage like Oil-Red O. Oil-Red O has been validated to be the most effective way to stain triglyceride mass in *C. elegans*[14].

The final assay described here is the nuclei staining of *C. elegans* mitotic region in the germline. Briefly, germ cells are in the mitotic phase located at the distal tip of the germline. This region contains germline stem cells as well. As these mitotic cells leave the mitotic region and travel proximally towards the gonads, they enter the meiotic cell cycle[15]. The goal in staining for mitotic cells is to understand if an increase in proliferation of 16p11.2 deletion orthologous genes gives rise to an increase/decrease in body area (Fig. 4). This phenotype is comparable to a clinical presentation of macrocephaly or microcephaly, respectively.

Human Gene	*C. elegans* ortholog	*C. elegans* Ortholog Description
MAPK3	*mpk-1*	Mitogen-activated protein kinase, functions in worm vulval cell fate specification, cell migration/guidance, defense against bacterial infection
ALDOA	*aldo-2*	Fructose-bisphosphate aldolase, active in the elongation of the worm embryo
ASPDH-1	*K09A9.6*	Expressed in body wall musculature, vulval muscle, ventral nerve cord, and the head and tail neuron.
DOC2A	*rbf-1*	Sole *C. elegans* rabphilin homolog; required for normal basal rates of locomotion
KCTD13	*sdz-35*	Ortholog of human potassium channel tetramerization domain, tumor necrosis factor, alpha-induced protein 1
CORO1A	*cor-1*	Actin filament binding protein homolog coronin, aids in cytoskeleton organization.
TAOK2	*kin-18*	Involved in feeding behavior, regulation of embryonic development, reproduction, nematode larval development and locomotion

Table 2. Selected orthologous genes within 16p11.2 deletion. For a complete list of phenotypes resulting from these genes, see Table 4.

C. elegans as a model for neurodevelopmental disorders

A *C. elegans* model of neurodevelopmental disorders provides a middle ground between the complexity of vertebrate models, such as mice, and the extreme simplicity of yeast. The first advantage of using a nematode model in the field of neuroscience is that is has a well-described neuronal lineage that provides a strong foundation for studying neuronal mechanisms. Secondly, the ability to genetically manipulate and identify genes for neuronal knockdown is made easy through RNAi. Third, a nematode is completely transparent, allowing for facile monitoring of neuronal or cellular components[16]. The hermaphrodite *C. elegans* has 302 neurons divided in 118 neuronal classes and 56 glial cells, altogether comprising 37% of all the somatic cells in the

worm. Neuronal classes include 39 classes of predicted sensory neurons, 27 of motor neurons, and the remainder as interneurons[17]. Considering its simplicity, *C. elegans* use all of the classical neurotransmitters such as acetylcholine, dopamine, serotonin, GABA and glutamate[16].

While we cannot ignore that nematodes do not emulate human behavior in its entirety, this animal model is still attractive in studying the genetics of neuronal phenomena. It is easy to genetically manipulate, feasible of doing a large-scale analysis, and has a quick reproductive cycle. Among all models, *C. elegans* is the most cost effective and still offers the benefit of a multicellular organism.

CHAPTER 2

MATERIALS AND METHODS

Worm Strains

The Caenorhabditis Genetics Center provided the TU3311 *Caenorhabditis elegans* strain, [unc-119p::YFP + unc-119p::sid-1], referred to as uIs60. The uIs60 genotype is hypersensitive to neuronal RNAi by feeding. It contains detectable YFP in neurons. All worms under RNAi feeding were maintained under 15-20 degrees[18]. All stocks were grown on Nematode Growth Medium (NGM) agar[19] seeded with wild type OP50 bacteria and RNAi knockdowns were grown on RNAi plates[20]. Pre-cloned *C. elegans* genomic fragments in an L4440 plasmid into HT115 (DE3), an RNase III-deficient *E. coli strain,* were used from the Ahringer library to induce RNA interference[21]. The following Ahringer RNAi clones were used to target homologous genomic sequences of the 16p11.2 deletion in worms: III-2I07; X-7M01; III-3G21; III-3F13; II-9J09; III-3O13; III-5D01; and II-3E05. RNAi feeding assays for both single hits and double hits, through double feeding, were carried out as described[22]. Constructs were validated by sequencing.

Egg Synchronization

The ages of worms were normalized for all assays by synchronizing the hatching of eggs. The bleaching technique was used for egg synchronization as previously described[23]. Adult worms (mixed stages) were collected off NGM plates with 3 ml of M9 buffer and centrifuged at 1000 x g for 1 minute. The supernatant was discarded and the pellet was treated with 400µl of 100% bleach and 150µl of 3M NaOH. The solution was inverted 6 times and then incubated for 3 minutes. The mixture was diluted with 10 ml of M9 buffer. The tube was inverted and the

mixture was centrifuged at 4000 x g for 1 minute. The eggs were washed 3 times with 10 ml of M9 buffer. Eggs were rocked overnight and incubated at 20°C.

Thrashing Assay

C. elegans eggs were synchronized and grown up on RNAi clone-expressing HT115 bacteria. At the L4 stage, between 20-30 worms movements were recorded on a Photometrics® CoolSNAP EZ camera for 10 seconds after adding 1 ml of M9 buffer. Thrashing activity was quantified by the ImageJ plugin "wrMTrck." Movements of all worms were thresholded to 70% area change and the bendThreshold was held constant at 2.0. MinSize, maxSize, maxVelocity, and minTrackLength parameters were also held constant at 1000 $pixel^2$, 20000 $pixels^2$, 50 pixels/frame , and 10 frames, respectively.

Germline Proliferation Assay

Animals were synchronized and grown up on RNAi clone-expressing HT115 bacteria. At the L4 stage, worms were transferred into a 2 µl LB drop on a glass etched ring slide. The worms were treated with a large drop of Carnoy's fixative. Slide was allowed to dry for 1 minute. About 3-5µl of Slowfade Antifade 4,6-diamidoino-2-phenylindole (DAPI) was added to the center of the etched ring. Between 5-10 animals were imaged by confocal fluorescent microscopy on a ZEISS LSM 800 with Airyscan. Imaging was focused on the germline, specifically the distal tip region. ImageJ was used to quantify the width of the gonad.

Oil-Red O Staining Assay

Oil Red O staining was performed in reference to O'Rourke et. al.[14] Per genotype, 2-3 plates of synchronized L4 animals were collected with 1 x PBS. Worms were washed three times with 1x PBS pH 7.4 and allowed to settle by gravity. In order to permeabilize the cuticle, 120µl of PBS and 120 µl of 2X MRWB containing 2% paraformaldehyde was added to the pellet (2x MRWB buffer: 160 mM KCl, 40 mM NaCl, 14 mM Na2EGTA, 1 mM spermidine-HCl, 0.4 mM spermine, 30 mM Na-PIPES pH 7.4, 0.2% ß-mercaptoethanol). The mixture was gently rocked and incubated at room temperature for 1 hour. Animals equilibrated with 60% isopropanol for 15 minutes. After allowing the worms to settle, isopropanol was removed and worms were incubated overnight at room temperature in 1 mL of 60% Oil-Red O stain while rocking. Dye was removed, and 200 µl of 1x PBS .01% Triton X-100 was added. Animals were mounted on agarose slides and imaged on an Olympus DP73 color digital microscope camera, outfitted with brightfield and widefield optics.

Oil-Red O ImageJ Analysis

Processing of images and thresholding was performed using ImageJ. The intensity of the Oil-Red O stain was measured by calculating Mean Gray Intensity values. Normalized 8-bit images were thresholded, using a constant threshold, so that only the highest 20% of the Oil Red O signals were recorded. Therefore, the background stain from inhomogeneous structures was eliminated and only the darkest stain corresponding to the intestinal tract was measured. Mean Gray Intensity, Minimum Gray Intensity, and Median Gray Intensity values were calculated through the Set Measurements tool in ImageJ. These values were compared for each genotype to the EV control using an unpaired t test.

Expression Profiling

Gene expression of 7 *C. elegans* genes, homologous to genes encoded within the human 16p11.2 deletion, were confirmed through qRT-PCR. Primer design, validation, qRT-PCR conditions and data analysis were performed as previously described. Normalization was performed using 2 reference-genes; *tba-1* (tubulin alpha chain) and *cdc-42* (cell division control protein 42).

Human Gene	Sequence
***mpk-1* FWD**	CGAGGAAGTTCATGGGCAAC
***mpk-1* REV**	AGATTCACATAACGGGGAGCA
***aldo-2* FWD**	GTCACCCCAGGAAAGGGAAT
***aldo-2* REV**	CATGGATCCGGTGGACTCATC
***K09A9.6* FWD**	AAATCCGGCGTTCCGACG
***K09A9.6* REV**	AGAACCGGATTGACGCCAG
***cdc-42* FWD**	TGCCGACAGTCTTCGACAAT
***cdc-42* REV**	GCTCGCCACCGATCATTACT
***tba-1* FWD**	ACACTCCACTGATCTCTGCTG
***tba-1* REV**	GACAGAGAGAGCCTCATGGTAAG
***kin-18* FWD**	CCGTTTTTCATGGCCCCAG
***kin-18* REV**	TAGCCGTGGTTTTCATTTCCA
***rbf-1* FWD**	CTGGACCACAGCCAATGACA
***rbf-1* REV**	ATGAGCATGAGTATGCCGGG
***sdz-35* FWD**	CGGGTTTTTCAAAACACTGCTG
***sdz-35* REV**	AGAAGCAGTAGGTCCATGACG
***cor-1* FWD**	AACCAAAGGTCCAGGTAGCC
***cor-1* REV**	CACGTGCCCTACGTCAGA

Table 3. Selected orthologous gene primers. Used IDT primer design and BLAST to design primers and check targets.

CHAPTER 3

RESULTS

Single Hits Thrashing Assay

In order to understand the contribution of a single deletion on neuromuscular phenotypes when it is not coupled with another deletion hit, a thrashing, or swimming assay was carried out on animals containing a single 16p11.2 deletion gene. As previously described, the thrashing assay is a more demanding muscular assay compared to normal location on the agar plate (crawling). Animals with empty vector (EV) averaged at around 43.2 body bends per minute (BBPM). The negative control, *adsl-1*, was around 41.4 body bends per minute. Genotypes *mpk-1*, *K0A9.6*, *rbf-1*, *sdz-35*, and *cor-1* were all statistically insignificant to the empty vector control. Genotypes *aldo-2* and *kin-18* were statistically thrashing faster than the empty vector control (Fig. 2). *Aldo-2* and *kin-18* knockdowns were thrashing 25% and 30% faster than the empty vector control, respectively. The maximum body bends per minute recorded for *aldo-2* and *kin-18* were 72 BBPM for both.

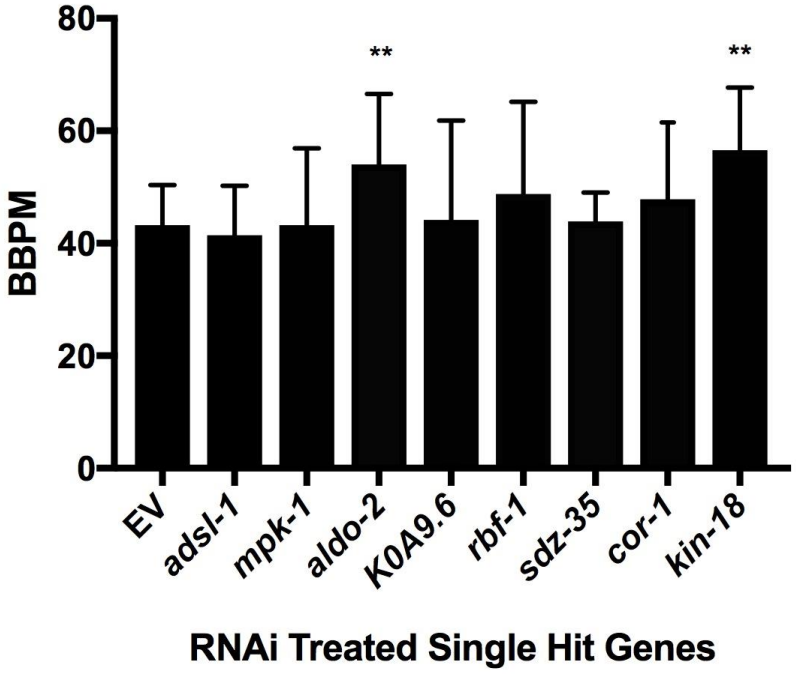

Figure 2. The 16p11.2 deletion single hits of orthologous *C. elegans* genes effects on neuromotor behavior. For each genotype, 15-20 animals were allowed to thrash on a standard NGM plate with 1 mL of M9 buffer. Videos were recorded for 10 sec for each animal and analyzed through ImageJ wrMTrck plugin.

Double Hits Thrashing Assay

As previously discussed, it has been shown that rare CNVs may be correlated to a more severe neurodevelopmental presentation when two or more are present in a given individual. This Two-Hit Theory for predicting neurodevelopmental variations was tested in a *C. elegans* thrashing assay. Phenotypes may indicate how gene knockdown combinations contribute to the overall 16p11.2 deletion clinical presentation.

In order to deliver a double knockdown of two genes, animals were double fed. In order to account for a potential diluted knockdown dosage, the negative control was fed with ½ *mpk-1* plasmid-containing *E. coli* and ½ empty vector containing *E. coli*.

The *mpk-1* control for double hits of thrashing was bending 62.6 BBPM. All gene combinations were statistically insignificant compared to EV + *mpk-1* control except the *mpk-1* + *kin-18* and *mpk-1* + *pisy-1*. These two-hit gene knockdowns showed to be 51.2 and 50.1 BBPM, respectively. *Mpk-1* + *kin-18* and *mpk-1* + *pisy-1* were thrashing 18.5% and 20% slower than the *mpk-1* + EV control (Fig. 3A).

The *K0A9.6* control for the double hits thrashed 71.3 BBPM. All gene combinations were statistically insignificant other than *K0A9.6* + *pisy-1* and *K0A9.6* + *kin-18*. *K0A9.6* + *pisy-1* thrashed 54.7 BBPM while *K0A9.6* + *kin-18* thrashed 41.7 BBPM. The *K0A9.6* + *pisy-1* knockdowns were thrashing 24.2% less than the *K0A9.6* controls. The *K0A9.6* + *kin-18* knockdown animals were thrashing ~30% less than that of controls (Fig. 3B).

The *aldo-2* control for the double hits thrashed on average 80.45 BBPM. Several gene combinations stood out in hindering the ability of animals to thrash. The two most significant gene combinations that hindered the ability of *C. elegans* individuals were *aldo-2* + *mpk-1* and *aldo-2* +*kin-18*. *Aldo-2* + *mpk-1* knockdowns thrashed 59.7 BBPM on average and 25.7% less

than controls, while the *aldo-2 + kin-18* knockdowns thrashed 49.3 BBPM on average and 31.07% less than that of the controls. Other gene combinations that were slightly statistically significant were *aldo-2 + sdz-35* and *aldo-2 + rbf-1*. *Aldo-2 + sdz-35* thrashed on average 68.6 BBPM, 11.85% less than that of controls. Aldo-2 + rbf-1 thrashed on average 69.41 BBPM, 11% less than that of controls. All other gene combinations were statistically insignificant (Fig. 3C).

The *cor-1* double hit control animals thrashed on average 62.5 BBPM. Interestingly enough, for *cor-1* double hit mutants, there were no gene combinations that showed a behavior that decreased average BBPM. Three double hit gene combinations showed statistically significant faster thrashing behavior. *Cor-1 + aldo-2*, *cor-1 + K0A9.6*, and *cor-1 + sdz-35* showed to statistically increase the thrashing behavior by 23.1%, 29.2%, and 23%, respectively. All other gene combinations were statistically insignificant (Fig. 3D).

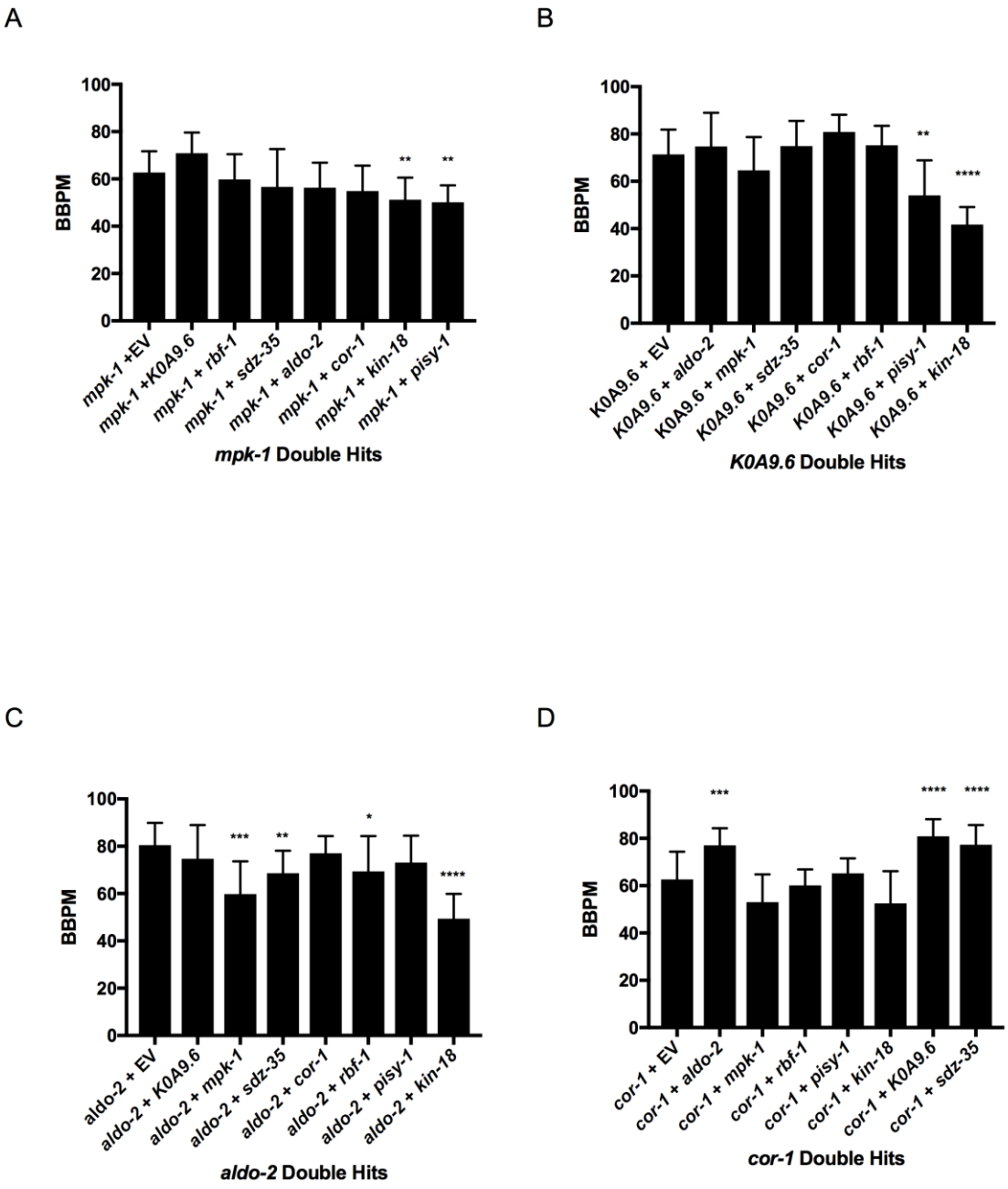

Figure 3. The Effect of Selected Double Hits of 16p11.2 Deletion Genes on Thrashing Neuromuscular Behavior. A. Animals were double fed *mpk-1* expressing HT115 bacteria in each combination. B. Animals were double fed *K0A9.6* expressing HT115 bacteria in each combination. C. Animals were double fed *aldo-2* expressing HT115 bacteria in each combination. D. Animals were double fed cor-1 expressing HT115 bacteria in each combination. For each genotype, 15-20 L4 animals were assayed for thrashing in a minute. Body bending analysis was completed via ImageJ wrMTrck plugin in order to disregard bias in manual counting.

Oil-Red O Intensities

The goal of staining for fat content on the animals was not to search for 16p11.2 deletion genes that could contribute to obesity. It has been shown previously that SH2B1 is a human gene already implicated in fat metabolism, and its mutant could potentially contribute to a congenital comorbidity of neurodevelopmental disease and obesity.

The goal of staining for fat content was to rule out the increase in body area in certain genotypes being due to increased fat content. If a knockdown has an increased body area, but insignificant body fat, it has the potential to have increased cellular germline proliferation. Increased proliferation in the *C. elegans* germline may be representative of increased head sizes seen in neurodevelopmental disorder presentations.

Oil-Red O has been widely accepted as the most effective method to stain triglyceride mass in *C. elegans*. Nile Red and BODIPY-labeled fatty acids are other lipid stains that were widely accepted to dye fat content, until Nile Red was shown to increase in intensity as fasting time increased. Additionally, BODIPY-labeled fatty acids were shown to fail at staining *C. elegans* major fat storage compartments. It was shown to stain lysosome related organelles, not fat storage along the intestinal tract like Oil- Red O.

Three measurements were taken per animal: Mean Gray Intensity Value, Median Gray Intensity Value, and Minimum Gray Intensity Value. Mean Gray Intensity Value was calculated through ImageJ by converting an RGB image to an 8-bit grayscale image. On a scale of 0 to 255, 0 indicative of black and 255 indicative of white, each pixel in the worm was given a corresponding gray value.

In order to accurately measure intensity corresponding to the region of fat content on the worm, the intestinal tract, the gray values of all pixels within the worm were thresholded to 20%.

This ImageJ threshold has been previously published as a standard for measuring fat on the worm in an anatomically correct location. In this way, background coloration is excluded from analysis. The Mean Gray Intensity Value (Fig. 5) was calculated by taking the mean of the top 20% darkest area pixels. Additionally, to compare how different the darkest regions are between each single hit, the Minimum Gray Intensity Value was calculated (Fig 6). Two positive controls *sbp-1* (sterol regulatory binding protein 1) and *nhr-49* (nuclear hormone receptor 49) were used to represent a low and high fat phenotype, respectively. *Sbp-1* dramatically reduces the fat content and biogenesis of intestinal lipid droplets. Targets include acyl-CoA carboxylase and fatty acid synthase. *Nhr-49* has been shown to down-regulate mitochondrial beta-oxidation genes and causes high fat accumulation.

All 7 homologous genes were tested for potential increase in body fat, however only data for *mpk-1* knockdowns are shown. Consistently across more than 15 *mpk-1* animals, the mean gray intensity showed no change compared to the EV control. However, the body area (Fig. 4) was significantly increased in every life stage (L1 through adult). The empty vector control, on average had a gray pixel intensity of 101.469 while the *mpk-1* knockdowns had a mean gray pixel intensity of 87.08 (Fig. 5).

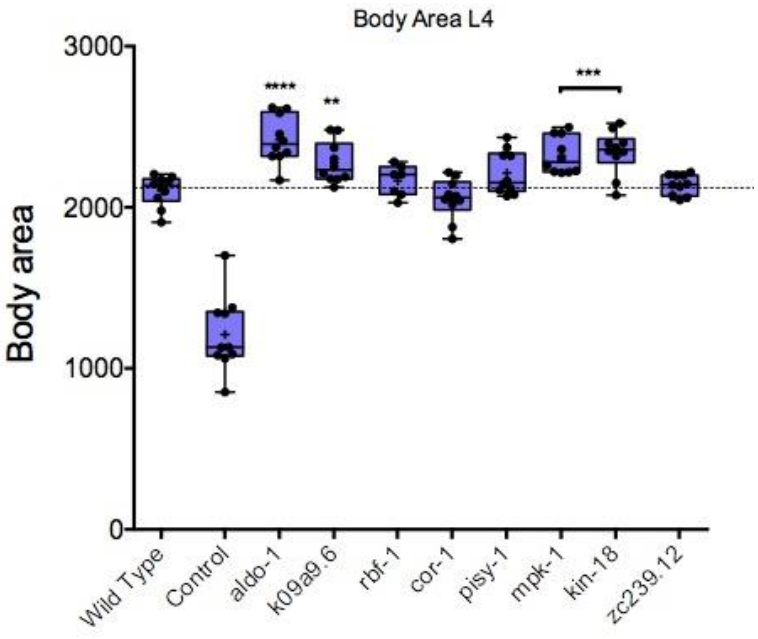

Figure 4. Body areas of all RNAi-treated single hits. An increase in *mpk-1* compared to wildtype is observed. 8-bit images analyzed with ImageJ. Credit: Ayush Thomas

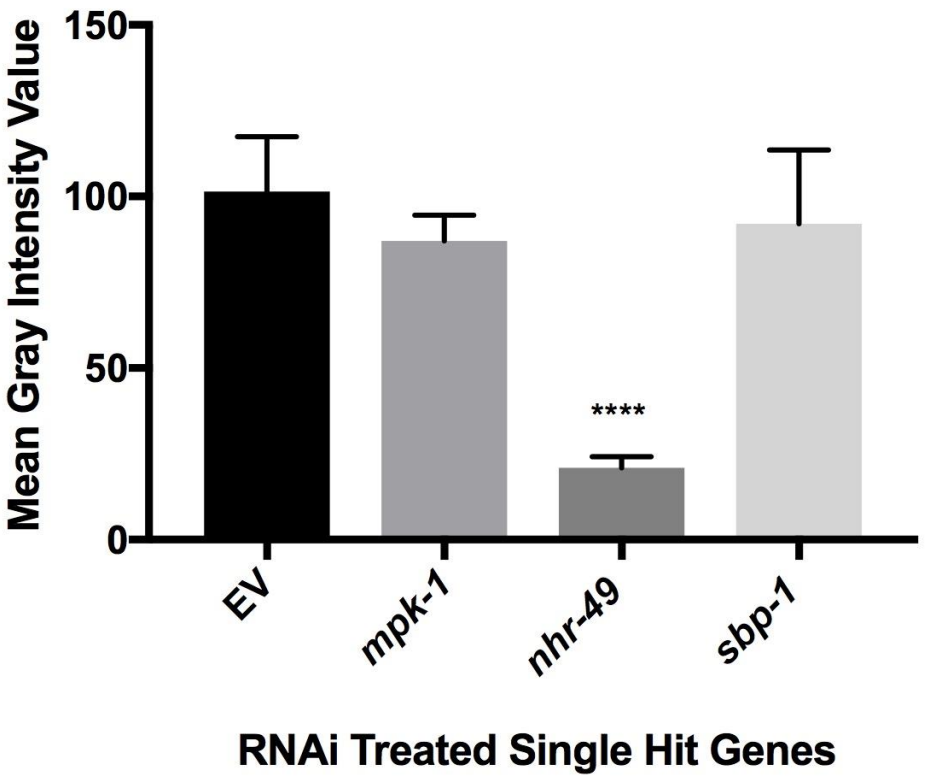

Figure 5. *mpk-1* knockdowns show no change in amount of fat, but increased body area. Controls for Oil-Red O staining were *nhr-49* (nuclear hormone receptor) and *sbp-1* (sterol regulatory binding protein 1). Approximately 20-30 animals were analyzed per genotype. *Mpk-1* knockdowns consistently had an insignificant amount of fat within the intestinal region compared to EV control.

Additionally, to compare the intensity of the darkest stained pixels between EV and *mpk-1* knockdowns, the Minimum Gray Intensity Value (corresponding to a low intensity value – closer to black at a score of zero) was calculated through the ImageJ 20% thresholding method. The minimum (darkest) intensity value across 15 *mpk-1* knockdown animals was consistently higher (not significantly) than that of controls (Fig. 6). Figure 7 shows a visual comparison of *mpk-1* knockdown animals with negative and positive controls.

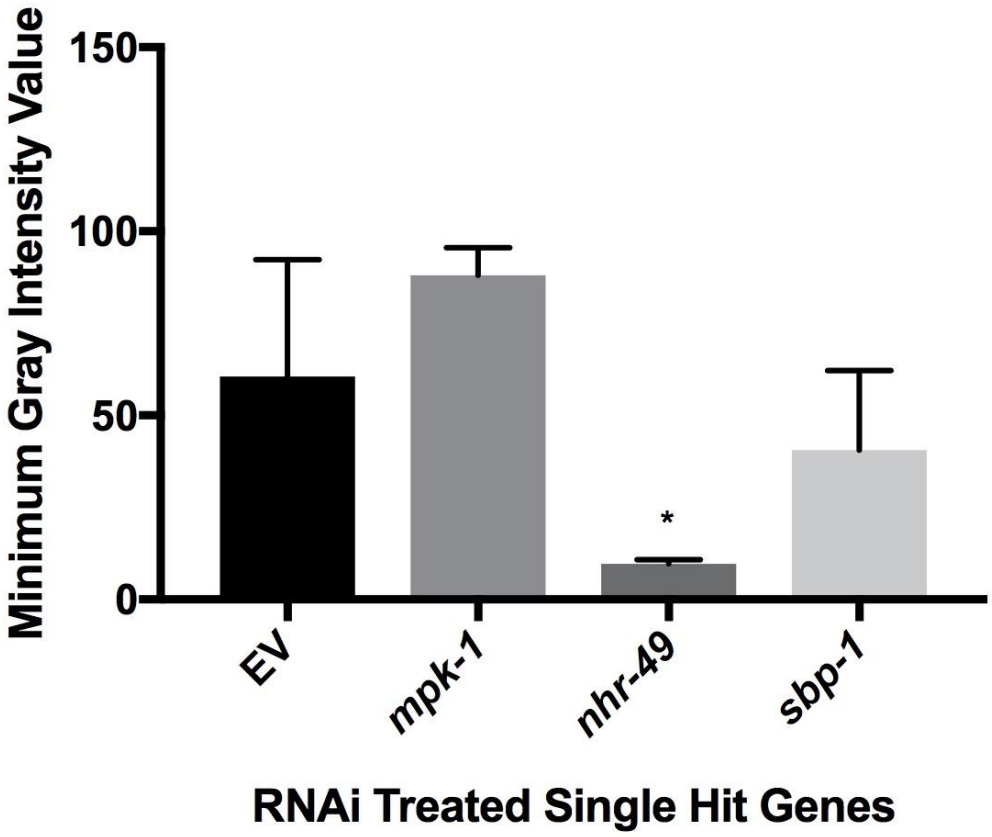

Figure 6. The darkest regions on *mpk-1* knockdowns are lighter than those of controls. Though not significant, there was still a lower minimum intensity for *mpk-1* animals compared to controls. The average minimum intensity values across 20-30 animals for EV controls was 60.5 compared to that of *mpk-1* knockdowns, which was 88.

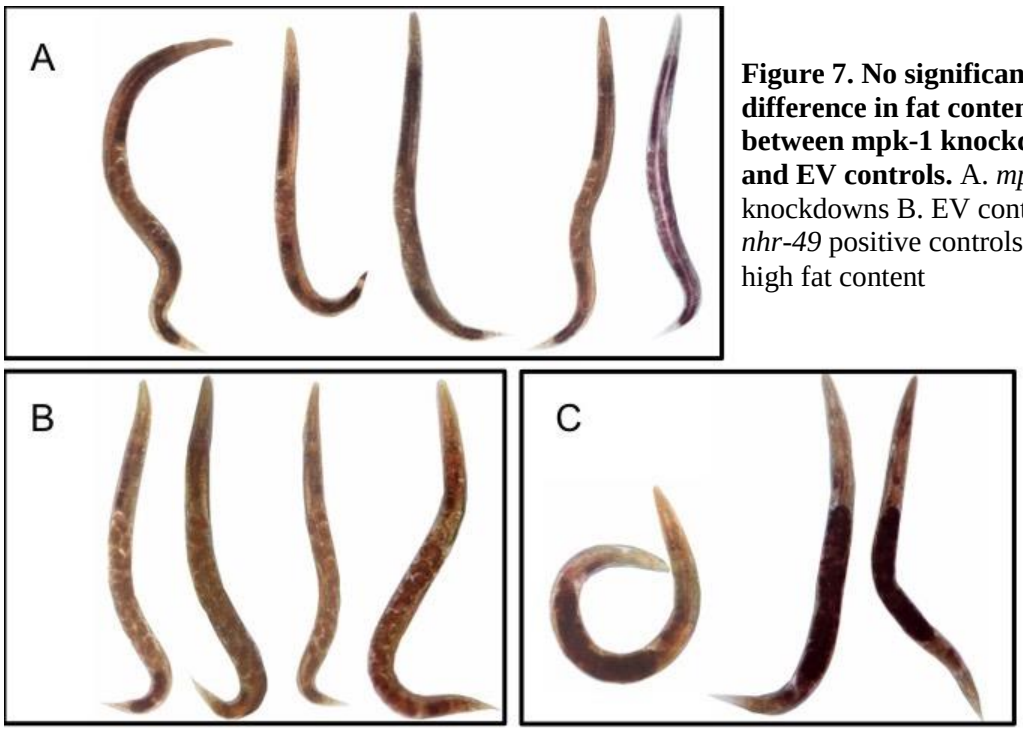

Figure 7. No significant difference in fat content between mpk-1 knockdown and EV controls. A. *mpk-1* knockdowns B. EV controls C. *nhr-49* positive controls showing high fat content

Germline Proliferation in *mpk-1* knockdowns

When the body area of *mpk-1* RNAi knockdown worms are measured, the body area is always significantly increased when compared to controls (Fig. 4). In order to assess the causation of an increased body size, and Oil Red O staining assay was done in order to rule out the effects of increase lipid content. As mentioned above, there was consistently no difference in the lipid content of *mpk-1* knockdowns.

The next step was to test if an increase in proliferation in the mitotic region of the gonad

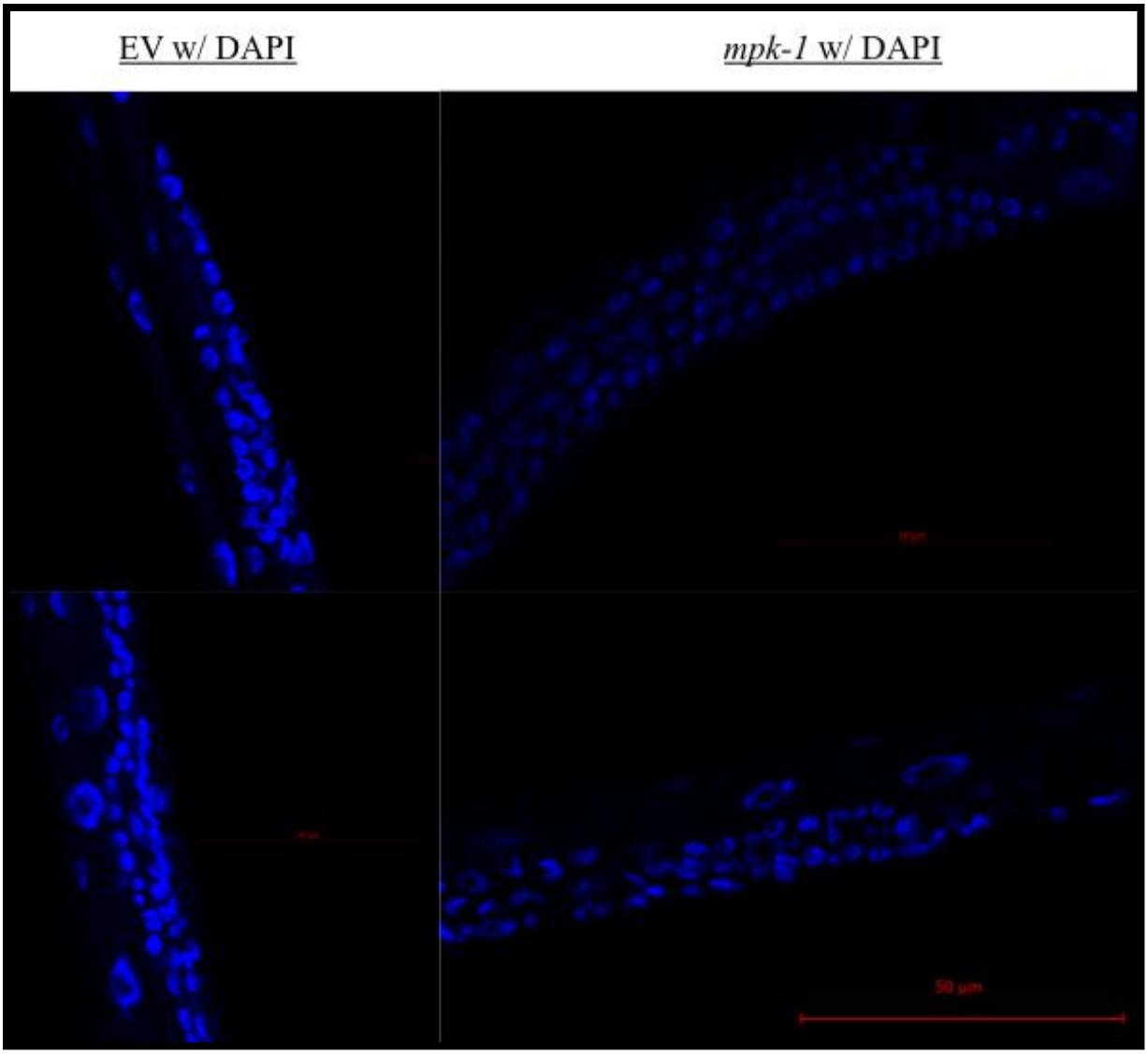

Figure 8. DAPI stained nuclei in distal tip of gonad. A gonad arm extrusion procedure was not used in order to preserve the integrity of the gonad; however, the gonad can be easily spotted by identifying the region of highly ordered cells in the mitotic region.

could be the cause. There is a visible increase in the width of the gonad just by looking at the stained nuclei that make up the gonad arm. The mitotic region of the gonad arm can easily be spotted by identifying a region of highly ordered cells that stem from a common point, which is the distal tip cell (DTC).

The number of cells present in the gonad were counted from the DTC to a length of 100μm. There was both a significant increase in cellular proliferation (Fig. 9A) and width of gonad (Fig. 9B) in *mpk-1* knockdowns versus controls. It should be noted that an increase in proliferation of the C. elegans germline would not affect the length of the worm; the gonad arm bends back toward the midline in order for the mitotic cells to be fertilized. Only an increase in the width of the gonad arm would increase the size of the worm.

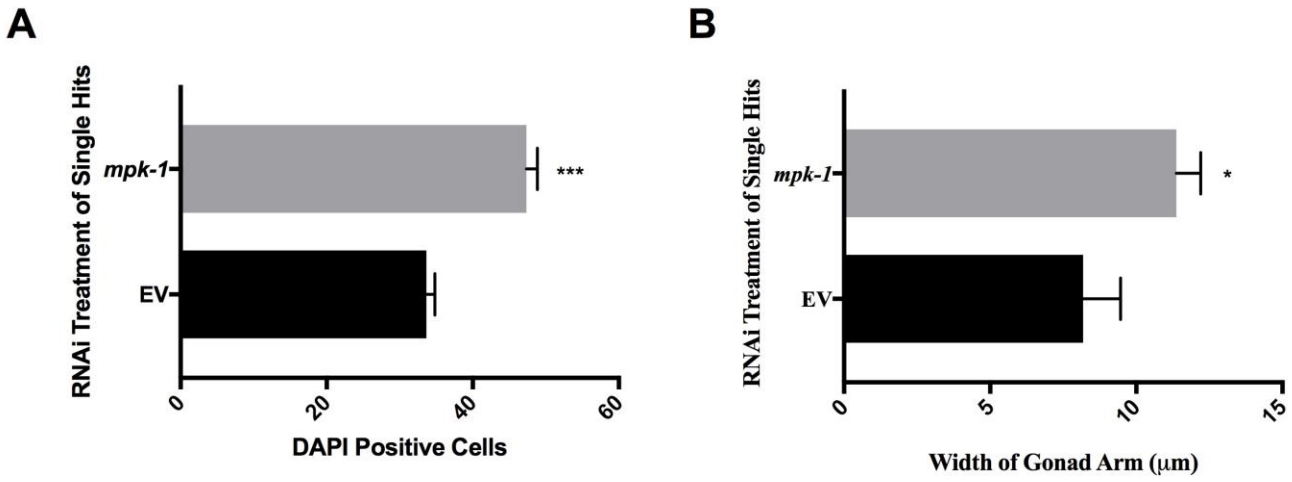

Figure 9. An increase in proliferation gives rise to an overall larger animal. Significant increase in body area of *mpk-1* knockdowns compared to EV not shown. A. Results from a DAPI positive cell count within the first 100μm of the DTC show both a significant increase in the number of cells and B. a significant increase in the width of the gonad.

CHAPTER 4

DISCUSSION

In order to understand the individual gene contributions to the human 16p11.2 recurrent microdeletion, RNAi was used to knockdown orthologous genes in *C. elegans*. Understanding these individual gene contributions impacts how physicians and scientists could potentially identify, diagnose, and treat patients. Recapitulating a standard phenotype of hyperactivity, motor dysfunction, or macrocephaly could help pin-point which genes could be targeted for gene therapy in human patients.

The thrashing locomotion assay was used to screen for a potential phenotype of motor dysfunction. In summary, the RNAi treatment of single hits resulted in *aldo-2* and *kin-18* having an increased number of body bends per minute (Fig. 2). Not surprisingly, both *aldo-2* and *kin-18* are both involved in elongation of the embryo and locomotion, respectively (Table 2). However, it may at first seem unintuitive that a knockdown of these genes would cause an increase in body area. I hypothesize that there may be an overcompensation in another gene when each of these are knocked out individually. It's possible that these two genes overcompensate for each other. When *aldo-2* and *kin-18* are both knocked down in the same animal, there is a significant decrease in the number of body bends per minute (Fig. 10). This significant decrease is the second highest decrease seen in all of the knockdowns.

Additionally, two other gene combinations in the thrashing double hits see a significant decrease when in combination with *aldo-2*. Both *aldo-2 + sdz-35* and *aldo-2 + mpk-1* show a significant decrease in the number of body bends were minute. *Aldo-2*, when knocked down alone, sees an increase in body bends per minute, perhaps due to another gene overcompensating for its absence. However, when coupled with other genes, a significant decrease in the number of

body bends per minute is seen. This phenotype is indicative of neuromotor defects, recapitulating a canonical 16p11.2 recurrent microdeletion presentation.

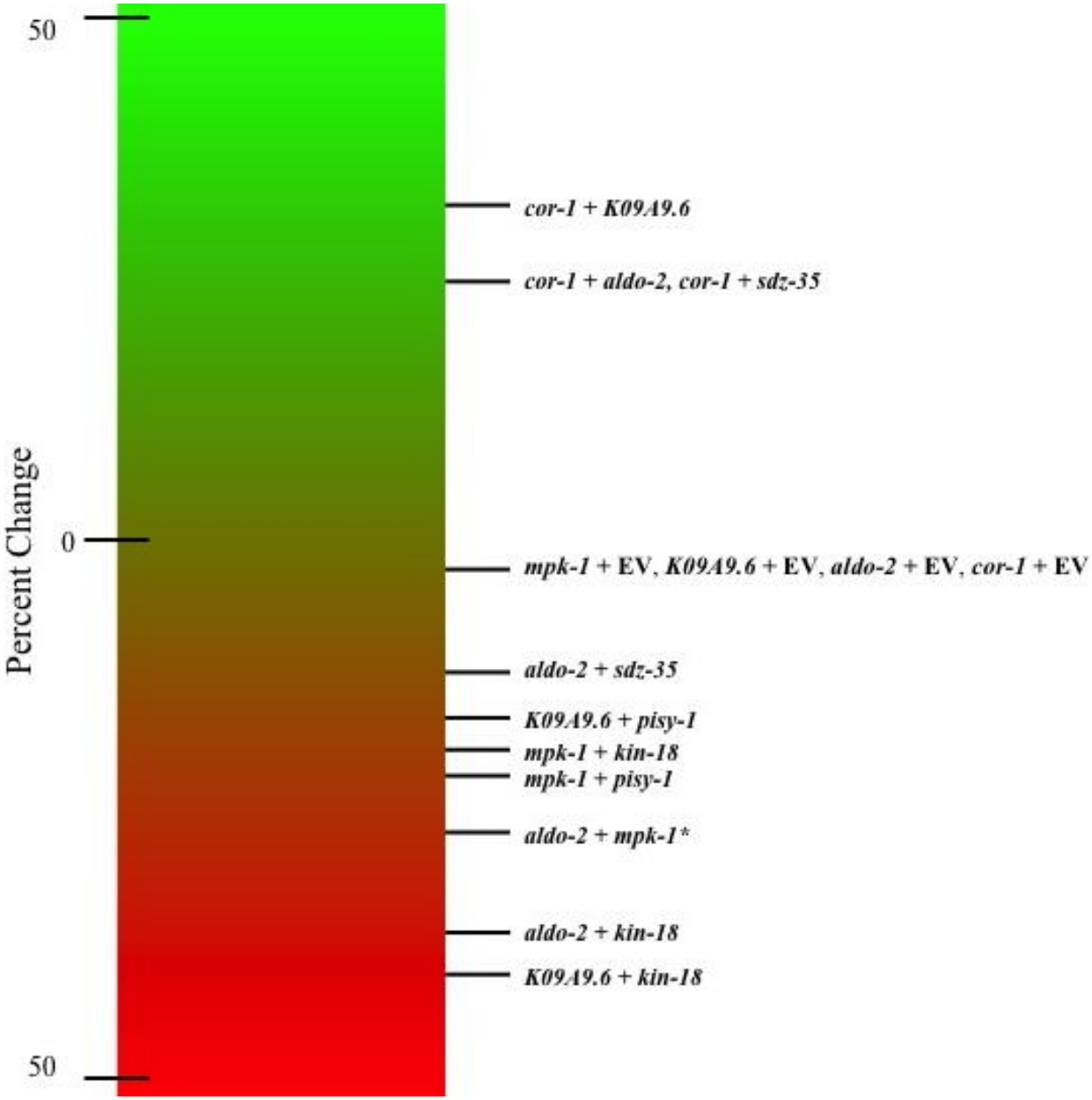

Figure 10. Summary of RNAi Treated Double Hits for Thrashing Assay. Each gene combination is compared to its respective control. Green is an increase in the number of body bends per minute while red shows a decrease in the number of body bends per minute. **aldo-2 + mpk-1* knockdowns here are compared to the *aldo-2* control.

One result that was not expected was the double-hit results with *cor-1*. *Cor-1* is crucially involved in the cytoskeleton movement and cell locomotion. Thrashing single hit results show no difference from the control. However, when combined with *K09A9.6*, *sdz-35*, and *aldo-2*, a significant increase in the number of body bends is seen. It is unclear if these genes are interacting in a way that is causing an increase in body bends per minute. However, based on unpublished data in our lab, *cor-1* knockdowns have been shown to be significantly smaller in size (length) than controls (Fig. 4). It is possible that the smaller length of worm is confounding the ImageJ ability to accurately track its bendThreshold. One may hypothesize that a shorter length of worm would require less energy to bend at the same angle of which a very large worm would bend. A smaller size may "rescue" the lack of *cor-1* gene.

Another recurring phenotype associated with the human 16p11.2 deletion is the increase in head size, otherwise known as macrocephaly. This *C. elegans* model has elucidated 1 gene that may give rise to this phenotype: *mpk-1*. In order to first screen for which orthologous gene could potentially contribute to this phenotype, the body areas of every single-hit gene was measured (Fig. 4). Of those genes, *mpk-1* was significantly larger in body size ($pixel^2$) than controls. The next step was to screen for a potential increase in fat content on the animals. Additionally, obesity is correlated with the 16p11.2 deletion, so it would be logical if one of the genes were involved in a higher fat phenotype.

After quantification of the Oil Red-O staining with ImageJ, there was no significant increase in body fat for *mpk-1* knockdowns. This was surprising, since *mpk-1* knockdown animals were consistently larger than controls.

Staining for increased germline proliferation could point to why there is an increase in body area. After staining and quantifying the number of DAPI positive cells in the mitotic region

of the gonad, a significant increase was seen in *mpk-1* knockdowns. This resulted is validated by other peer reviewed publications[15]. *Mpk-1* has been shown to function in promoting the proliferative fate and membrane organization.

In this study, *mpk-1* was identified as a player in determination the overall body size due to increased cellular proliferation. Additionally, several gene combinations and single hits were identified as players in a neuromotor dysfunction phenotype. Elucidating the individual contributions of these genes towards the overall 16p11.2 deletion phenotype helps physicians change the diagnosis paradigm, from that of a behavior evaluation process to that of a genomic analysis.

Table 4. Gross phenotypes arising from RNAi treatment of selected 16p11.2 deletion orthologs

Human Gene	Worm Sequence ID	Worm Gene	RNAi clone	Phenotype Detected	Thrash	ORO	Proliferation
MAPK3	F43C1.2	*mpk-1*	III-2I07				increased
ASPDH-1	K09A9.6	*K09A9.6*	X-7M01			lighter	
ALDOA	F01F1.12	*aldo-2*	III-3G21	✓	faster		
DOC2A	F37A4.7	*rbf-1*	III-3F13				
CDIPT	Y46G5A.5	*pisy-1*	II-9J09				
TAOK2	T17E9.1	*kin-18*	III-3O13	✓	faster		
CORO1A	R01H10.3	*cor-1*	III-5D01			lighter	
KCTD13	ZC239.12	*sdz-35*	II-3E05				
MAPK3 + ASPDH-1	*	*mpk-1 + K09A9.6*	*	✓	faster		
MAPK3 +ALDOA	*	*mpk-1 + aldo-2*	*	✓	slower		
MAPK3 + DOC2A	*	*mpk-1 + rbf-1*	*				
MAPK3 + CDIPT	*	*mpk-1 + pisy-1*	*	✓	slower		
MAPK3 + TAOK2	*	*mpk-1 + kin-18*	*	✓	slower		
MAPK3 + CORO1A	*	*mpk-1 + cor-1*	*				
MAPK3 + KCTD13	*	*mpk-1 + sdz-35*	*				
ASPDH-1 + ALDOA	*	*K09A9.6 + aldo-2*	*				
ASPDH-1 + DOC2A	*	*K09A9.6 + rbf-1*	*				
ASPDH-1 + CDIPT	*	*K09A9.6 + pisy-1*	*	✓	slower		
ASPDH-1 + TAOK2	*	*K09A9.6 + kin-18*	*	✓	slower		
ASPDH-1 +CORO1A	*	*K09A9.6 + cor-1*	*	✓	faster		
ASPDH-1 + KCTD13	*	*K09A9.6 + sdz-35*	*				
ALDOA + DOC2A	*	*aldo-2 + rbf-1*	*	✓	slower		
ALDOA + CDIPT	*	*aldo-2 + pisy-1*	*				
ALDOA + TAOK2	*	*aldo-2 + kin-18*	*				
ALDOA + CORO1A	*	*aldo-2 + cor-1*	*	✓	faster		
ALDOA + KCTD13	*	*aldo-2 + sdz-35*	*	✓	slower		
DOC2A + CDIPT	*	*rbf-1 + pisy-1*	*				
DOC2A + TAOK2	*	*rbf-1 + kin-18*	*				
DOC2A + CORO1A	*	*rbf-1 + cor-1*	*				
DOC2A + KCTD13	*	*rbf-1 + sdz-35*	*				
CDIPT + TAOK2	*	*pisy-1 + kin-18*	*				
CDIPT + CORO1A	*	*pisy-1 + cor-1*	*				
CDIPT + KCTD13	*	*pisy-1 + sdz-35*	*				
TAOK2 + CORO1A	*	*kin-18 + cor-1*	*				
TAOK2 + KCDTD13	*	*kin-18 + sdz-35*	*				
CORO1A + KCDT13	*	*cor-1 + sdz-35*	*	✓	faster		

References

1. Neurodevelopmental Disorders. In: *Diagnostic and Statistical Manual of Mental Disorders*. American Psychiatric Association; 2013. doi:10.1176/appi.books.9780890425596.dsm01
2. Acsadi G. *New Advances in Pediatric Neurologic and Developmental Disorders in the Era of Genomics*. Elsevier Health Sciences; 2015. https://books.google.com/books?id=6fvLCgAAQBAJ&dq=New+Advances+in+Pediatric+Neurologic+and+Developmental+Disorders+in+the+Era+of+Genomics,+An+Issue+of+Pediatric+Clinics+of+North+America,+E-Book&source=gbs_navlinks_s. Accessed March 2, 2018.
3. Hu WF, Chahrour MH, Walsh CA. The Diverse Genetic Landscape of Neurodevelopmental Disorders. *Annu Rev Genomics Hum Genet*. 2014;15(1):195-213. doi:10.1146/annurev-genom-090413-025600
4. Torres F, Barbosa M, Maciel P. Recurrent copy number variations as risk factors for neurodevelopmental disorders: critical overview and analysis of clinical implications. *J Med Genet*. 2016;53(2):73-90. doi:10.1136/jmedgenet-2015-103366
5. Redon R, Ishikawa S, Fitch KR, et al. Global variation in copy number in the human genome. *Nature*. 2006;444(7118):444-454. doi:10.1038/nature05329
6. Cooper GM, Coe BP, Girirajan S, et al. A copy number variation morbidity map of developmental delay. *Nat Genet*. 2011;43(9):838-846. doi:10.1038/ng.909
7. Buckingham SD, Sattelle DB. Strategies for automated analysis of C. elegans locomotion. *Invertebr Neurosci*. 2008;8(3):121-131. doi:10.1007/s10158-008-0077-3
8. Hempel M, Brugués NR, Wagenstaller J, et al. Microdeletion syndrome 16p11.2-p12.2:

Clinical and molecular characterization. *Am J Med Genet Part A*. 2009;149(10). doi:10.1002/ajmg.a.33042

9. Drake J, Link CD, Butterfield DA. Oxidative stress precedes fibrillar deposition of Alzheimer's disease amyloid beta-peptide (1-42) in a transgenic Caenorhabditis elegans model. *Neurobiol Aging*. 24(3):415-420. http://www.ncbi.nlm.nih.gov/pubmed/12600717. Accessed March 2, 2018.
10. Alexander AG, Marfil V, Li C. Use of Caenorhabditis elegans as a model to study Alzheimer's disease and other neurodegenerative diseases. *Front Genet*. 2014;5:279. doi:10.3389/fgene.2014.00279
11. Burt EC, Towers PR, Sattelle DB. Caenorhabditis elegans in the study of SMN-interacting proteins: a role for SMI-1, an orthologue of human Gemin2 and the identification of novel components of the SMN complex. *Invert Neurosci*. 2006;6(4):145-159. doi:10.1007/s10158-006-0027-x
12. van Ham TJ, Thijssen KL, Breitling R, Hofstra RMW, Plasterk RHA, Nollen EAA. C. elegans model identifies genetic modifiers of alpha-synuclein inclusion formation during aging. Kim SK, ed. *PLoS Genet*. 2008;4(3):e1000027. doi:10.1371/journal.pgen.1000027
13. Mullaney BC, Ashrafi K. C. elegans fat storage and metabolic regulation. *Biochim Biophys Acta*. 2009;1791(6):474-478. doi:10.1016/j.bbalip.2008.12.013
14. O'Rourke EJ, Soukas AA, Carr CE, Ruvkun G. C. elegans Major Fats Are Stored in Vesicles Distinct from Lysosome-Related Organelles. *Cell Metab*. 2009;10(5):430-435. doi:10.1016/J.CMET.2009.10.002
15. Hajnal A, Berset T, Lamont LB, Wickens M, Kimble J. The C.elegans MAPK phosphatase LIP-1 is required for the G(2)/M meiotic arrest of developing oocytes. *EMBO*

J. 2002;21(16):4317-4326. doi:10.1093/emboj/cdf430

16. Bessa C, Maciel P, Rodrigues AJ. Using C. elegans to Decipher the Cellular and Molecular Mechanisms Underlying Neurodevelopmental Disorders. *Mol Neurobiol*. 2013;48(3):465-489. doi:10.1007/s12035-013-8434-6
17. Bono M de, Villu Maricq A. NEURONAL SUBSTRATES OF COMPLEX BEHAVIORS IN C. ELEGANS. *Annu Rev Neurosci*. 2005;28(1):451-501. doi:10.1146/annurev.neuro.27.070203.144259
18. Calixto A, Chelur D, Topalidou I, Chen X, Chalfie M. Enhanced neuronal RNAi in C. elegans using SID-1. *Nat Methods*. 2010;7(7):554-559. doi:10.1038/nmeth.1463
19. Brenner S. The genetics of Caenorhabditis elegans. *Genetics*. 1974;77(1):71-94. http://www.ncbi.nlm.nih.gov/pubmed/4366476. Accessed March 2, 2018.
20. Timmons L, Court DL, Fire A. Ingestion of bacterially expressed dsRNAs can produce specific and potent genetic interference in Caenorhabditis elegans. *Gene*. 2001;263(1-2):103-112. http://www.ncbi.nlm.nih.gov/pubmed/11223248. Accessed March 2, 2018.
21. Fraser AG, Kamath RS, Zipperlen P, Martinez-Campos M, Sohrmann M, Ahringer J. Functional genomic analysis of C. elegans chromosome I by systematic RNA interference. *Nature*. 2000;408(6810):325-330. doi:10.1038/35042517
22. Ahringer J. Reverse genetics. *WormBook*. 2006. doi:10.1895/wormbook.1.47.1
23. Porta-de-la-Riva M, Fontrodona L, Villanueva A, Cerón J. Basic Caenorhabditis elegans methods: synchronization and observation. *J Vis Exp*. 2012;(64):e4019. doi:10.3791/4019

Academic Vita

EDUCATION

Penn State Medical-Scientist Training Program (MSTP) **Hershey, PA**
Medical Doctor & Doctorate of Philosophy *Anticipated Graduation in May 2026*

- Matriculating into M1 year in July 2018
- Ph.D. Track in Neuroscience

Penn State Eberly College of Science **University Park, PA**
Bachelor of Science in Biochemistry & Molecular Biology *Graduation May 2018*

- Concentration: Cellular Biology
- Schreyer Honors College Scholar
- Presidential Leadership Academy (PLA) Certificate
- Resident Assistant in Sorority Special Living Option

PUBLICATIONS & POSTERS

***Caernorhabditis elegans* model for 16p11.2 recurrent microdeletion suggests differential effects of conserved genes to knockdown**.
Anticipated Submission Date May 2018

- Sullivan, R. Thomas, A. Manuscript ready for submission to PLOS Genetics in May 2018.

Poster: Elucidating obesity and Autism Spectrum Disorder (ASD) comorbidity with a novel *Caenorhabditis elegans* 16p11.2 deletion model. Oct. 2016

- 2016 University Park Research Exhibition

Poster: Characterization of novel Wntless interacting proteins to understand the molecular mechanisms of addiction Aug. 2016

- 2016 MD/PhD Summer Exposure Program Poster Exhibition
- Penn State Hershey Medical Center
- Mentored by Dr. Robert Levenson and Dr. Jessica Petko

Identification of novel Wntless interacting proteins to understand the molecular mechanisms of opioid addiction: The search for how Wntless helps you "want" less. Aug. 2015

- 2015 MD/PhD Summer Exposure Program Poster Exhibition
- Penn State Hershey Medical Center
- Mentored by Dr. Robert Levenson and Dr. Jessica Petko

RESEARCH EXPERIENCE

Girirajan Lab at The Pennsylvania State University Sept. 2014 – May 2018

- Studying the behavioral and morphological phenotypes of *C. elegans* with deletions in the 16p11.2 locus.
- Worked with confocal fluorescent microscopy, metabolomics, and a nematode model system

Levenson Lab at the Hershey Center for Applied Research May 2015 – Aug. 2016

- Identified protein interactors of Wls to further investigate its relationship with the mu-opioid receptor
- Worked with a yeast model system
- Leveraged standard molecular biology techniques

HONORS & AWARDS

Student Engagement Network Grant (SEN) May 2017

- Awarded over $2000 to buy microscopes for the students of Pueblo Del Sol High School in Phoenix, AZ.
- Used grant money to teach underserved minority students about research as a career path and how to use microscopes.
- Developed and taught a one-week science class at Camp Catanese STEM camp

2016 Summer Erickson Discovery Research Grant May 2016

Goldwater Scholarship Nominee December 2015

Penn State BMB SURF Award April 2015

- Declined to accept Penn State Hershey MD/PhD Exposure Internship