THE PENNSYLVANIA STATE UNIVERSITY
SCHREYER HONORS COLLEGE

DEPARTMENT OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

CAERNORHABDITIS ELEGANS MODEL FOR 16P11.2 RECURRENT MICRODELETION SUGGESTS DIFFERENTIAL EFFECTS OF CONSERVED GENES TO KNOCKDOWN

AYUSH THOMAS
SPRING 2018

A thesis
submitted in partial fulfillment
of the requirements
for a 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

David Gilmour
Professor of Molecular and Cell Biology
Honors Adviser

* Signatures are on file in the Schreyer Honors College.

ABSTRACT

The 16p11.2 deletion syndrome is a rare (<1% of population) copy number variation (CNV) associated with a wide range of neurodevelopmental disorders, with a prevalence of 0.6% of patients for Autism Spectrum Disorder (ASD), 0.4% of patients with intellectual disability, and developmental birth defects.[1,2,3] Such neurodevelopmental disorders are characterized by diminished cognitive functionality and motor defects. The extensive genetic diversity associated with neurodevelopmental disorders causes their underlying functionality to remain unknown. In this study, one-hit and two-hit knockdown of orthologous 16p11.2 genes in *C. elegans* models elucidated the cellular and behavioral phenotypes and trends associated with each individual gene. We tested eight neurodevelopmental genes to document differential sensitivities of these genes in dosage alterations. Of the eight orthologous genes present between *C. elegans* and *H. Sapiens*, motor defects and body area phenotypes were commonly associated with the suppression of *aldo-1*, *k09a9.6*, and *mpk-1* human orthologs. Following the neurodevelopmental two-hit model proposed by Girirajan et. al., two-hit knockdowns of the eight orthologous genes were paired systematically, and revealed that two-hit pairings with *aldo-1*, *k09a9.6*, and *mpk-1* demonstrated increased sensitivity in neurodevelopmental and behavioral phenotypes in relation to one-hit models. These results suggest the importance of further studies of the mechanistic factors associated with morphological changes and motor dysfunction within the 16p11.2 region.

---

[1] Bijlsma EK et al. (2009) Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. Eur J Med Genet 52:77-87.

[2] Rosenfeld JA, Coppinger, J., Bejjani, B.A., Girirajan, S., Eichler, E.E., Shaffer, L.G., Ballif., B.C. (2010) Speech delays and behavioral problems are the predominant features in individuals with developmental delays and 16p11.2 microdeletions and microduplications. J Neurodevelop Disord 2:26-38.

[3]Shinawi M, Liu P, Kang SH, Shen J, Belmont JW, Scott DA, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. Journal of medical genetics. 2010;47(5):332–41. Epub 2009/11/17.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

First and foremost, I would like to give my appreciation for the support and advice Dr. Santhosh Girirajan had given me throughout my time in the Girirajan lab. His relentless pursuit in ensuring that each of his students reach their full potential was inspiring and had helped me tremendously in the process of the developing of my research study and the writing of my thesis. I want to thank Dr. David Gilmour for providing additional guidance throughout the entire process and for being a supportive Honors Advisor. I would also like to give a special thanks to Lucilla Pizzo and Rhea Sullivan, for without them, this project would not have been successful. A special acknowledgement is made to the Hanna-Rose Lab for allowing the use of their laboratory equipment and sharing their reservoir of *C. elegans* knowledge. I appreciate all the help they have provided over the last four years.

I am also thankful for the support from grants provided by the Eberly College of Science and the Schreyer Honors College. These grants provided me with the confidence and opportunities necessary for successful completion of my thesis. Finally, a huge thank you goes to my family and friends for their support through the writing of my thesis.

Chapter 1 : Introduction

An overview of past Autism Spectrum Disorders (ASDs) studies highlights the little history of genetic research in this field, as well as the assortment of the methods used. Prior to the 1970s, autism was not widely considered to have a significant biological basis. Instead, various psychodynamic interpretations, including the role of a cold or aloof style of mothering, were invoked as potential causes. It was only in the 1980s that the significance of genetic influences became well-defined, when the co-occurrence of chromosomal disorders and rare syndromes associated with ASDs were recorded.[4] Technical advances in the 1990's encouraged the first candidate gene association studies, which was followed by whole-genome linkage studies that identified additional loci of potential interest. Through these studies, a large number of potentially important copy number variations (CNVs) were identified as potential candidates for ASDs phenotypes.[5,6] In contrast to the 1970s, it is now understood that defined mutations, genetic syndromes and *de novo* CNV account for about 10–20% of ASD cases.[7] In search of a genetically susceptible loci, several studies have discovered a recurrent ~600-kb microdeletion at 16p11.2 in multiple affected individuals. [8] It has been observed that CNVs associated with the

---

[4] Blomquist HK, et al. Frequency of the fragile X syndrome in infantile autism. A Swedish multicenter study. Clin. Genet. 1985;27:113–117.

[5] Sebat J, et al. Strong association of *de novo* copy number mutations with autism. Science. 2007;316:445–449. Beyond identifying important CNV likely to prove important to our understanding of the ASDs, this study highlights a significant difference in the frequency of *de novo* variants between simplex and multiplex families, raising the possibility that distinct mechanisms are involved in each.

[6] Jacquemont ML, et al. Array-based comparative genomic hybridisation identifies high frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. J. Med. Genet. 2006;43:843–849. The numerous *de novo* deletions that are reported in this work have received relatively little attention but are probably important in the ASDs. These data also show that rare *de novo* mutations are likely to appear at particularly high frequencies in syndromic populations.

[7] Abrahams BS, Geschwind DH. Advances in autism genetics: on the threshold of a new neurobiology. Nat Rev Genet. 2008;9:341–55. doi: 10.1038/nrg2346.

[8] Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, et al. Recurrent 16p11.2 microdeletions in autism. Hum Mol Genet. 2008;17:628–38. doi: 10.1093/hmg/ddm376.

16p11.2 microdeletion are associated with a wide range of neurodevelopmental outcomes such as Autism, hyperactivity, developmental delay, epilepsy, and obesity. [9] Overall, the 16p11.2 deletion accounts for 1% of sporadic autism cases.[10]

The model organism used for this neurodevelopmental study of the 16p11.2 deletion is the nematode worm, *Caenorhabditis elegans*. *Caernorhabditis elegans* is widely used in neuroscience thanks to its well-understood development and lineage of the nervous system. *C. elegans* species possesses 70% conserved counterparts to human genes, including eight genes associated with the human 16p11.2 chromosomal region. Using ensembl data bases, the eight human genes that were identified to overlap with the *C. elegans* organism are *aldo-1, k09a9.6, cor-1, rbf-1, mpk-1, pisy-1, kin-18,* and *zc239.12.* The orthologous *C. elegans* genes, along with each respective protein function are listed in table 1.

[9] Rosenfeld, J. A., Coppinger, J., Bejjani, B. A., Girirajan, S., Eichler, E. E., Shaffer, L. G., & Ballif, B. C. (2010). Speech delays and behavioral problems are the predominant features in individuals with developmental delays and 16p11.2 microdeletions and microduplications. *Journal of Neurodevelopmental Disorders*, *2*(1), 26–38.
[10] Mefford HC et al. (2010) Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. PLoS Genet 6:e1000962.

Nematode Ortholog	*H. sapiens* Gene	Protein	Function
aldo-1	*aldoa*	Fructose-bisphosphate aldolase 1	Active in the elongation of the worm embryo[11]
k09a9.6	*aspdh*	Uncharacterized Protein	Unknown Function
cor-1	*coro1a*	Coronin-like protein cor-1	Actin filament binding protein homolog coronin, aids in cytoskeleton organization[12]
rbf-1	*doc2a*	Rabphilin-1	Sole C. *elegans* rabphilin homolog; required for normal basal rates of locomotion[13]
mpk-1	*mapk3*	Mitogen-activated protein kinase	Functions in worm vulval cell fate specification, cell migration/guidance, cellular proliferation, defense against bacterial infection[14]
pisy-1	*cdipt*	CDP-diacylglycerol--inositol 3-phosphatidyltransferase	Lipid biosynthesis, lipid metabolism; Active in human and worm embryonic motor neurons
kin-18	*taok2*	Serine/threonine-protein kinase SULU	Invovled in feeding behavior, regulation of embryonic development, reproduction, nematode larval development and locomotion[15]
zc239.12	*kctd13*	Ortholog of human potassium channel tetramerization domain	Unknown function

Table 1: Human genes implicated in the 1611.2 deletion and corresponding C. *elegans* orthologous genes

To gain an understanding of the phenotypes associated with the respective genes listed in table 1, individual knockdowns were performed in the *C. elegans* model to observe for deviance from the wild-type population. The principal of performing genetic knockouts within *C. elegans*

[11] Inoue T, Yatsuki H, Kusakabe T, Joh K, Takasaki Y, Nikoh N, Miyata T, Hori K. Caenorhabditis elegans has two isozymic forms, CE-1 and CE-2, of fructose-1,6-bisphosphate aldolase which are encoded by different genes. Arch Biochem Biophys. 1997 Mar 1;339(1):226-34. PubMed PMID: 9056253.

[12] Wang X, Zhou F, Lv S, Yi P, Zhu Z, Yang Y, Feng G, Li W, Ou G. Transmembrane protein MIG-13 links the Wnt signaling and Hox genes to the cell polarity in neuronal migration. Proc Natl Acad Sci U S A. 2013 Jul 2;110(27):11175-80. doi: 10.1073/pnas.1301849110. Epub 2013 Jun 19. PubMed PMID: 23784779; PubMed Central PMCID: PMC3703986.

[13] Staunton J, Ganetzky B, Nonet ML. Rabphilin potentiates soluble N-ethylmaleimide sensitive factor attachment protein receptor function independently of rab3. J Neurosci. 2001 Dec 1;21(23):9255-64. PubMed PMID: 11717359.

[14] Schouest KR, Kurasawa Y, Furuta T, Hisamoto N, Matsumoto K, Schumacher JM. The germinal center kinase GCK-1 is a negative regulator of MAP kinase activation and apoptosis in the C. elegans germline. PLoS One. 2009 Oct 14;4(10):e7450. doi: 10.1371/journal.pone.0007450. PubMed PMID: 19826475; PubMed Central PMCID: PMC2757678.

[15] Berman KS, Hutchison M, Avery L, Cobb MH. kin-18, a C. elegans protein kinase involved in feeding. Gene. 2001 Nov 28;279(2):137-47. PubMed PMID: 11733138; PubMed Central PMCID: PMC4441751.

populations is to simulate reduced gene expression and induce genetic abnormalities within this region of interest. This was done in order to understand specific disease stimuli and conjunctive effects. Two assays were performed among knockout populations: Body Area Analysis and Crawling Speed Assays. Body size measurements were made from the analysis of flattened projections of the worms. This assay screened for abnormalities in the regulation of growth to a normal body size displayed among images of the wild-type control. Crawling speed rates measured locomotory behavior as *C. elegans* moved freely on agar plates, which allowed for unconstrained motion in order to encourage complex patterns for spontaneous motor behaviors. [16]

Investigations of the 16p11.2 chromosomal loci were additionally performed with *Drosophila* larval neuromuscular junctions to investigate the roles of these genes in relation to synaptic development. It was observed that the disruption of the human gene KIF22 and the gene product of Kinesin-2 complex caused ectopic innervations of axon branches, along with less frequent re-routing of axon branches.[17] Mouse models investigated that expression levels for the Major Vault Protein (MVP), a candidate gene in 16p11.2 microdeletion, as a key molecule influencing the homeostatic component of activity-dependent synaptic plasticity, and potentially the corresponding phenotypes of 16p11.2 microdeletion syndrome. Knockouts of 16p11.2 orthologs in Zebrafish suggested that a dysregulation of ceramide pathways and calcium sensitive exocytosis underlies seizures and large body sizes.[18]

---

[16] Croll N (1975) Components and patterns in the behavior of the nematode *Caenorhabditis elegans*. J Zool 176: 159–176.

[17] Park SM, Littleton JT, Park HR, Lee JH. Drosophila Homolog of Human KIF22 at the Autism-Linked 16p11.2 Loci Influences Synaptic Connectivity at Larval Neuromuscular Junctions. Exp Neurobiol. 2016 Feb;25(1):33-39.

[18] Jasmine M. McCammon, Alicia Blaker-Lee, Xiao Chen, Hazel Sive; The 16p11.2 homologs fam57ba and rbf-1 generate certain brain and body phenotypes, Human Molecular Genetics, Volume 26, Issue 19, 1 October 2017, Pages 3699–3712

C. elegans Structure

The *Caenorhabdidis elegans* is a small, soil-dwelling worm. The adult hermaphrodite is approximately 1 mm in length and 80 µm in diameter. Out of the 959 somatic cells, 302 are neurons connected by about 10,000 synapses.[19, 20] Although not fully characterized, the wiring diagram and the location of neuronal bodies have been well studied and found to be fairly reproducible from animal to animal. Despite the simplicity of the nervous system, the worm possesses a set of stereotypical behaviors: head movements, forward and backward locomotion, and turns.[21]

Figure 1: The *C. elegans* Model **An adult worm showing the pharynx (green), intestine (yellow), eggs (large blue ovals), neuronal cell bodies (small blue ovals), synapse-rich process bundles (red), commissural tracts (black). The major synapse-rich regions are the nerve ring (red surrounding the pharynx, the ventral nerve cord (red line with neuronal cell bodies on ventral side of animal), the dorsal nerve cord (red line on the dorsal side of the animal). Image reproduced from the *Nonet M.L. Laboratory Website.***

[19] White, J.G., E. Southgate, J.N. Thompson, and S. Brenner. 1986. The structure of the nervous system of the nematode Caenorhabditis elegans. Philos. Trans. R. Soc. Lond. B Biol. Sci. 314: 1-340.
[20] Durbin, R.M. 1987. Studies on the Development and Organisation of the Nervous System of Caenorhabditis elegans. In. Cambride University, Cambridge.
[21] Croll, N.A. 1975. Components and patterns in the behaviour of the nematode Caenorhabditis elegans. J. Zool. Lond 176: 159-176.

Due to their size, ease of cultivation, and short life cycle, we chose to work with these multi-cellular organisms to characterize the orthologous 16p11.2 genes and analyze nervous system development and morphological phenotypes.

RNAi Knockdown by Feeding

Systemic RNAi delivery can be accomplished by simply feeding dsRNA to *C. elegans* populations. Feeding RNAi is often the favored method due to its convenience and inexpensive form of dsRNA delivery. Pre-cloned *C. elegans* genomic fragments were constructed in an L4440 plasmid into HT115 (DE3). This RNase III-deficient *E. coli strain* was used from the Ahringer library to induce RNA interference. 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.

Hypothesis

As CNVs and deletions of regions within the 16p11.2 human chromosome are closely linked with phenotypes of Autism and ASDs, I predict that single-hit knockdowns of orthologous genes within *C. elegans* populations will induce a greater risk for motor defects and an increase of cellular proliferation. Additionally, I hypothesize that the introduction of two-hit knockdowns of these orthologous genes will cause a more pronounced phenotype due to the additional knockout of a functional gene.

Chapter 2 : Materials and Methods

Caenorhabditis elegans and E. coli

The Caenorhabditis Genetics Center (CGC) provided the TU3311 *Caenorhabditis elegans* strain, [unc-119p::YFP + unc-119p::sid-1], referred to as uIs60. The uIs60 genotype is a genetically engineered *C. elegans* strain with RNAi sensitive neurons. [22] All animals that were knocked down with RNAi were maintained at 20 degrees Celsius. All *uIs60* stocks were grown on Nematode Growth Medium (NGM) agar seeded with wild-type OP50 bacteria. RNAi feeding for experimentation was conducted on RNAi plates. RNAi constructs were verified by sequencing.

Culturing E. coli with RNAi

L4440 plasmid E. coli with desired gene inserts for knockdown were obtained from the Ahringer Library. 5mL of Lysogeny Broth (LB) with 5uL (100 µg/mL) ampicillin was added into eight sterilized 10mL glass tubes for each respective RNAi incorporated E. coli strain. Each culture solution was grown overnight and shaken at 37 °C. RNAi-expressing E. coli strains were preserved in a glycerol stock and stored within -80 ^{0}C freezer.

[22] Calixto A, Chelur D, Topalidou I, Chen X, Chalfie M. Enhanced neuronal RNAi in C. elegans using SID-1. Nature methods. 2010;7(7):554–9. pmid:20512143

E. coli RNAi Plates Preparation

RNAi incorporated E. coli strains obtained from the glycerol stocks were incubated overnight for approximately twelve hours in LB broth. 300 μL of each of the eight cultured RNAi-treated E. coli were pipetted onto an NGM plate for each respective trial. Two-hit gene knockdown experimentation was conducted using 150μL of each desired knockout RNAi incorporated E. coli strain, which brought the total volume to 300μL. This mixture was gently homogenized to allow for both RNAi-incorporated E. coli cultures to mix evenly. Each sample was pipetted onto RNAi plates and left to dry for 3 hours in preparation for seeding with L1 *C. elegans.*

Egg Synchronization

To obtain synchronized *C. elegans* for the purpose of each individual assay, adult egg-carrying *C. elegans* were raised on 10 NGM plates seeded with OP50 E. coli for 2-3 days. This allowed for mature *C. elegans* to cultivate on the plate. 2 mL of M9 buffer was pipetted onto each of the plates and swirled to dislodge the *C. elegans* from the medium. The M9/*C. elegans* mixture was collected into two 15 mL centrifuge tubes and spun down at 5000 x g. The supernatant containing the M9 solution was poured out leaving 250-400 μL of M9 solution with the *C. elegans* pellet. After gently re-suspending the *C. elegans*, lysis buffer (400 uL Bleach & 150 uL 3M NaOH) was added to each of the two 15mL centrifuge tubes and periodically inverted for 5 minutes until *C. elegans* were visibly dissolved and only eggs could be observed in the buffer solution. After the 5 minutes, 10 mL of M9 buffer solution was added to each of the 15 mL tubes to dilute and nullify the reaction. The 15 mL tubes were centrifuged at 9000 x g for 1 minute and the supernatant was poured out. M9 was added to the two tubes and centrifuged an additional two times within the

wash phase. After the last supernatant was dispensed, the remaining pellets were each suspended in 1.5 mL of M9 buffer. The two 15 mL centrifuge tubes were incubated on a shaker for 24 hours, which allowed the *C. elegans* eggs to grow and arrest at the L1 stage prior to seeding. 50 uL of M9/*C. elegans* L1 solution were plated the following day on fresh RNAi plates seeded with the respective RNAi incorporated E. coli.

Assay Setup

Individual assays were conducted in a series of trials using synchronized *C. elegans* that are arrested in the L1 phase. *C. elegans* were treated with the respective gene knockdown induced via oral-delivery of E. Coli RNAi plasmids. Treated animal populations were each monitored over each of the four growth phases (L1-L4) for phenotypic quantification. At the L4 growth phase, *C. elegans* populations were measured for crawling speed. Conditions for each trial were held under constant temperature (20 Celsius) and assays were performed in exact time intervals corresponding to the respective growth stages.

cDNA Preparation

After the knockout assay was conducted, RNA was extracted from the knockout populations to observe for effective knockdown. The sample was prepared to produce cDNA using the sample, water, and cDNA mix. The cDNA recovered was diluted 1:3, and was stored at -20C.

qPCR Analysis

mRNA expression was assessed using quantitative real-time PCR (qRT-PCR) by isolating RNA from *c. elegans* expressing RNAi knockdown of each specific genes. For RNA extraction, groups of 150–200 *C. elegans* were solubilized and isolated using TRIZOL (Invitrogen) and reverse transcribed using qScript cDNA synthesis kit (Quanta Biosciences). Quantitative RT-PCR was performed using an Applied Biosystems Fast 7500 system with SYBR Green PCR master mix (Quanta Biosciences). All SYBR green assays were performed in triplicate, and normalized to the mRNA expression of the housekeeping genes CDC-42, PMP-3, and TBA-1.

Chapter 3 : Results

Eight *C. elegans* orthologs of the human genes *pisy-1, k09a9.6-1, zc239.12, kin-18, rbf-1, aldo-1, mpk-1,* and *cor-1* were characterized using Ensembl online databases. Single-gene knockdowns were conducted to observe for phenotypic and behavioral trends upon gene suppression. Phenotypic assays included the measurement of body area across the L1 thru L4 growth stages, while the behavioral assays included measurement of the max and average crawling speed for each respective worm affected with the respective knockdown. Characterization of a single hit knockout was made in respect to the empty vector wild-type population. The positive control used for this assay was the knockout of the gene *adsl-1,* an adenylosuccinate lyase enzyme that attributes loss-of-function when suppressed. [23] Significance was determined using an unpaired student t-test, where P values measured to be less than 0.05 were determined as significant. Three trials of each single- hit knockdown was performed, in addition to verification of knock-down using RT-qPCR. Results and phenotypic trends from the single-hit assays can be seen in table 2.

[23] Chen, P., Wang, D., Chen, H., Zhou, Z., & He, X. (2016). The nonessentiality of essential genes in yeast provides therapeutic insights into a human disease. *Genome Research*, *26*(10), 1355–1362. http://doi.org/10.1101/gr.205955.116

	Worm Sequence ID	Human Gene	Fly Gene	Phenotype Detected	Body Area	Max Speed	Average Speed
aldo-1	WBGene00011474	aldoa	aldo-1				
k09a9.6	WBGene00185014	aspdh					
pisy-1	WBGene00012897	cdipt	pis				
cor-1	WBGene00000768	coro1a	cor-1				
rbf-1	WBGene00004316	doc2a	rph				
zc239.12		kctd13	CG10465				
mpk-1	WBGene00003401	mapk3					
kin-18	WBGene00002201	taok2					

Table 2: Phenotype map of single-hit knockdowns of 16p11.2 genes in *C. elegans* model **Red Spaces indicate that two out of three trials had data sets that possessed a p value of less than 0.05. Genes that are labeled as gray illustrate than an assay demonstrated a significant phenotype in at least two out of the three assays conducted.**

Body Area - Single Hits

Growth and measurement of body area was the first phenotype measured within our study of the eight genes associated within the 16p11.2 chromosomal region. Body area was measured by capturing images of individual *C. elegans* at the L1, L2, L3, and the L4 growth stage, and measuring approximately 10 individual worms from each stage by length and width.

Each specific *C. elegans* ortholog was knocked down upon the introduction of a gene-specific RNAi at the L1 stage via egg synchronization. *C. elegans* populations were monitored and imaged throughout their respective growth stages. Individual images were processed using the

software ImageJ. Significant trends were recorded in relation to deviance from the wild-type control, EV. Results were verified through a series of three trials, in which significant trends among genes were only recorded if trends matched two out of three trials. Body Area analysis of the eight single hit knockdowns from Trial 1 can be observed in Figure 2.

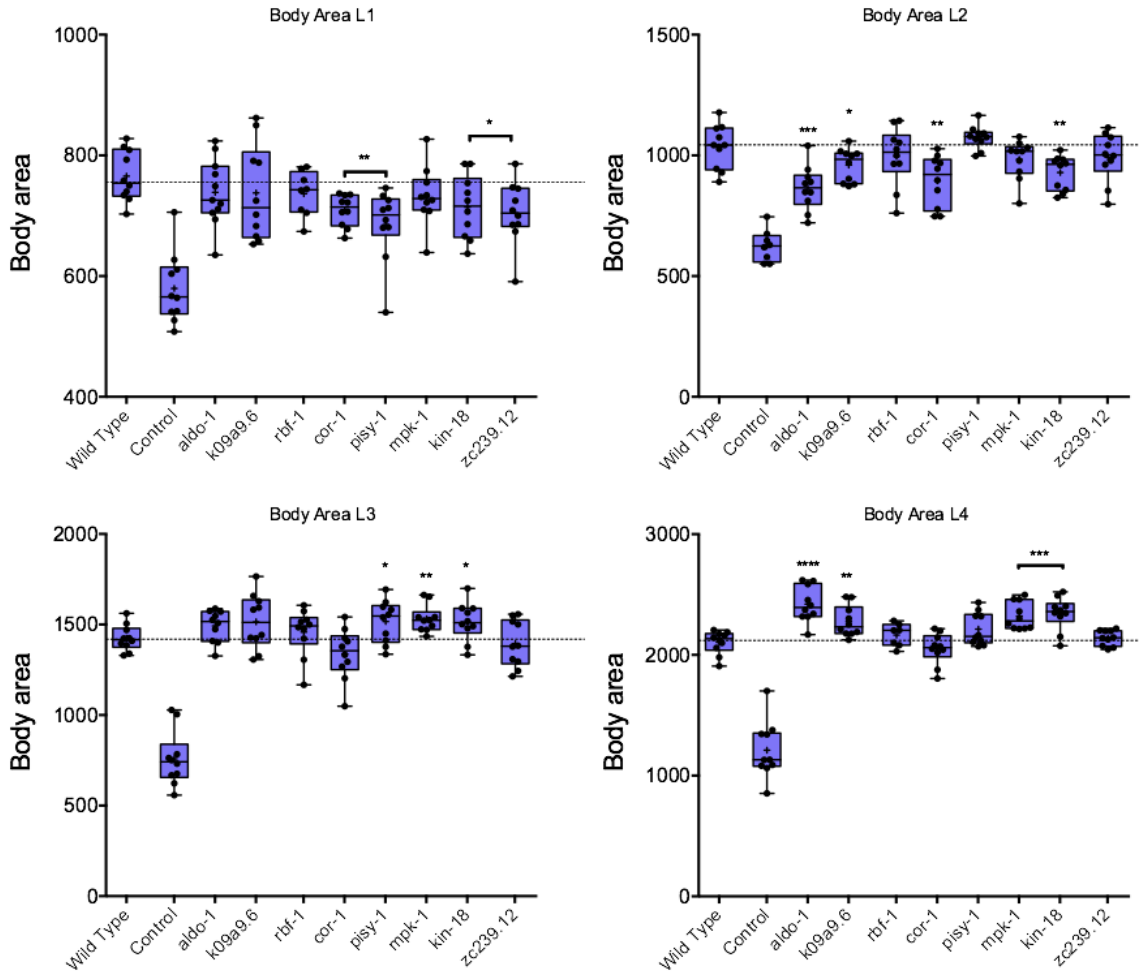


Figure 2: Trial 1 Body Area Analysis of One-Hit Knockdowns (N = 10) *C. elegans*. Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and ** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

Trends can first be observed within the L2 stage, where genes such as *aldo-1*, *k09a9.6*, *cor-1*, and *kin-18* were seen to be significantly smaller than the wild-type. At the L3 stage, the trend shifts towards worms increasing in size in relation to the wild-type, with *pisy-1*, *mpk-1*, and *kin-18* having significantly larger sizes. However, although not significant, it is notable that the knockout populations of *aldo-1* and *k09a9.6* are beginning to trend to larger worms in the L3 stage. At the L4 stage, *aldo-1*, *k09a9.6*, *mpk-1*, and *kin-18* continued the trend observed within the L3 stage, and were significantly larger than the wild-type population. Images of worms taken from the L4 stage are provided in Figure 3.

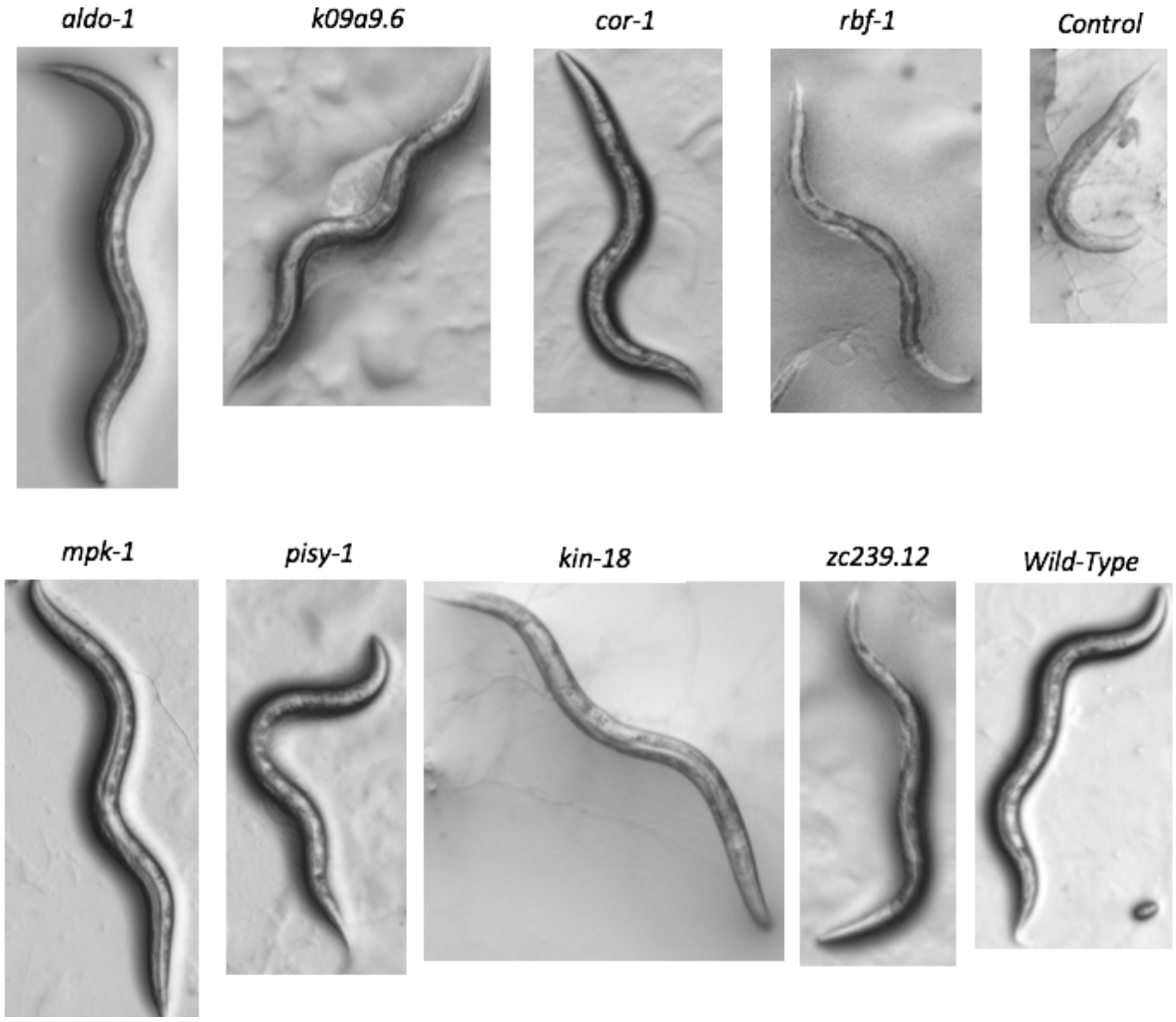


Figure 3: Trial 1 Images of 16p11.2 Gene Knockouts taken at the L4 Stage

Among the four growth stages, the greatest significance was seen during the L4 stage. Knockdowns of *aldo-1* ($p < 0.0001$), *k09a9.6* ($p < 0.01$), *mpk-1* ($p < 0.001$), and *kin-18* ($p < 0.001$) were significantly larger than the wild-type control in two out of the three trials performed. Trends associated with the *aldo-1*, *k09a9.6*, and *mpk-1* knockouts across the four growth stages was averaged and compared to the wild-type (EV) in Figure 4. The initial decrease of size within the L2 stage is observed for all three genes, prior to increasing in body area by the L3 and L4 stage in comparison to the Wild-type.

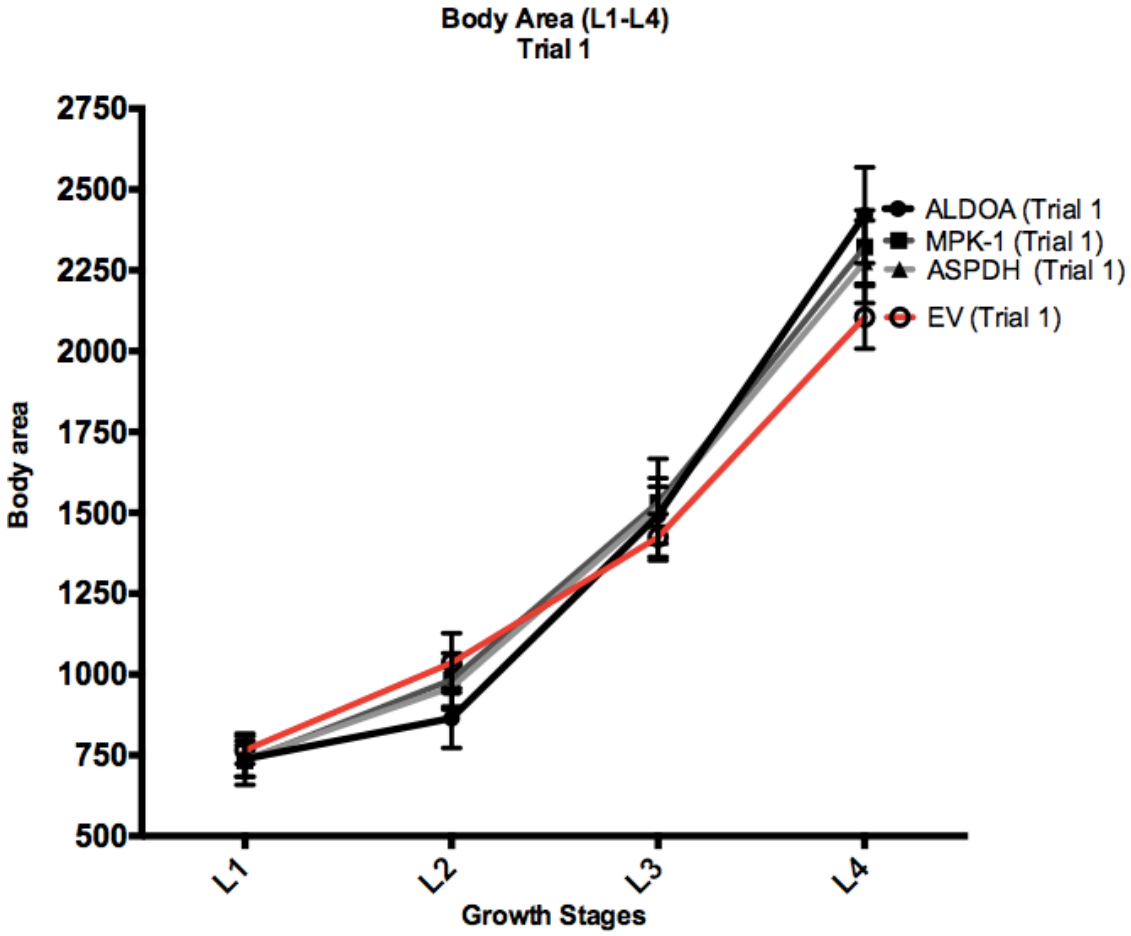


Figure 4: Trial 1 Body Area measurements across L1-L4 Growth Stages of aldo-1, mpk-1, k09a9.6, and wild-type (EV)

Crawling Speed – Single Hits

The average and max crawling speeds were recorded for individual *C. elegans* over a series of 10 seconds as they were in the L4 stage. The average speed of a worm tracked the average distance at which worms crawled over the measured time frame, while the max speed recorded the fastest rate achieved by individual worms over the given time frame. The average and max crawling speed of the eight knockout populations from trial 1 can be observed in Figure 5.

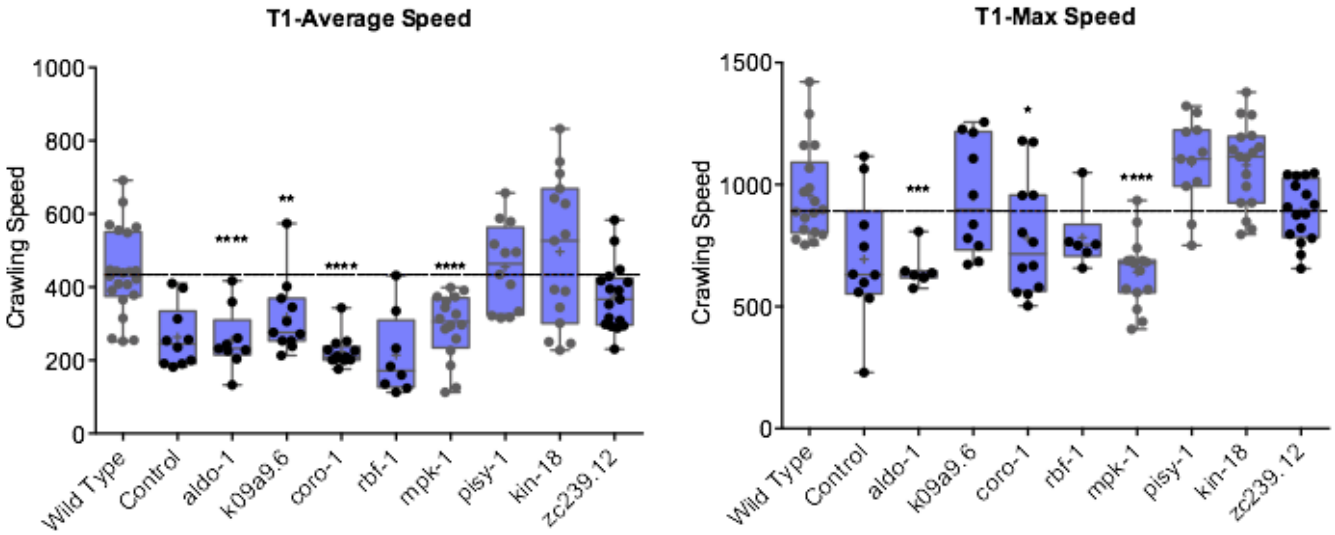


Figure 5: Trial 1 Average and Max speed measurements of 16p11.2 Gene knockouts **Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and **** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

Across three trials, an observed decrease in average speed was seen in knockouts of *aldo-1, k09a9.6-1, cor-1,* and *mpk-1.* The max speed followed a similar trend, as knockouts of *aldo-1, cor-1,* and *mpk-1* trended to lower depressed speeds in relation to the wild-type. Across these three assays, consistent genes that demonstrated phenotypes were *aldo-1, k09a9.6,* and *mpk-1.*

qPCR Verification – Single Hits

Single hit knockdowns were validated through qPCR Analysis. Gene expression was calculated in comparison to the *C. elegans* housekeeping gene, cdc-42. Genes that were successfully analyzed were *rbf-1, aldo-1, cor-1, k09a9.6,* and *mpk-1*. qPCR analysis of *pisy-1, kin-18,* and *zc239.12* were unsuccessful and were not included in the figure.

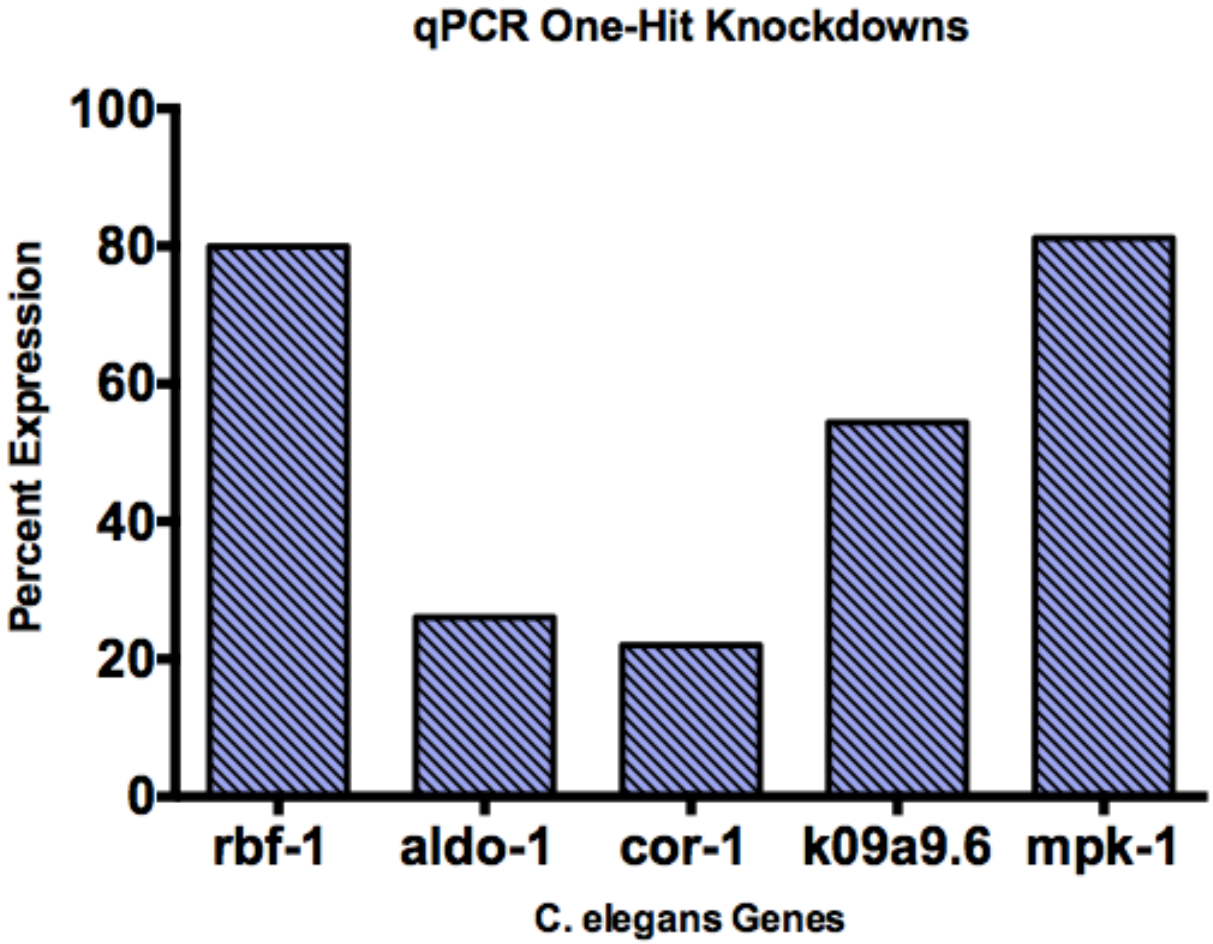


Figure 6: qPCR Verification of Single-Hit knock downs **Respective gene knockdowns were compared to the c. elegans housekeeping gene, cdc-42.**

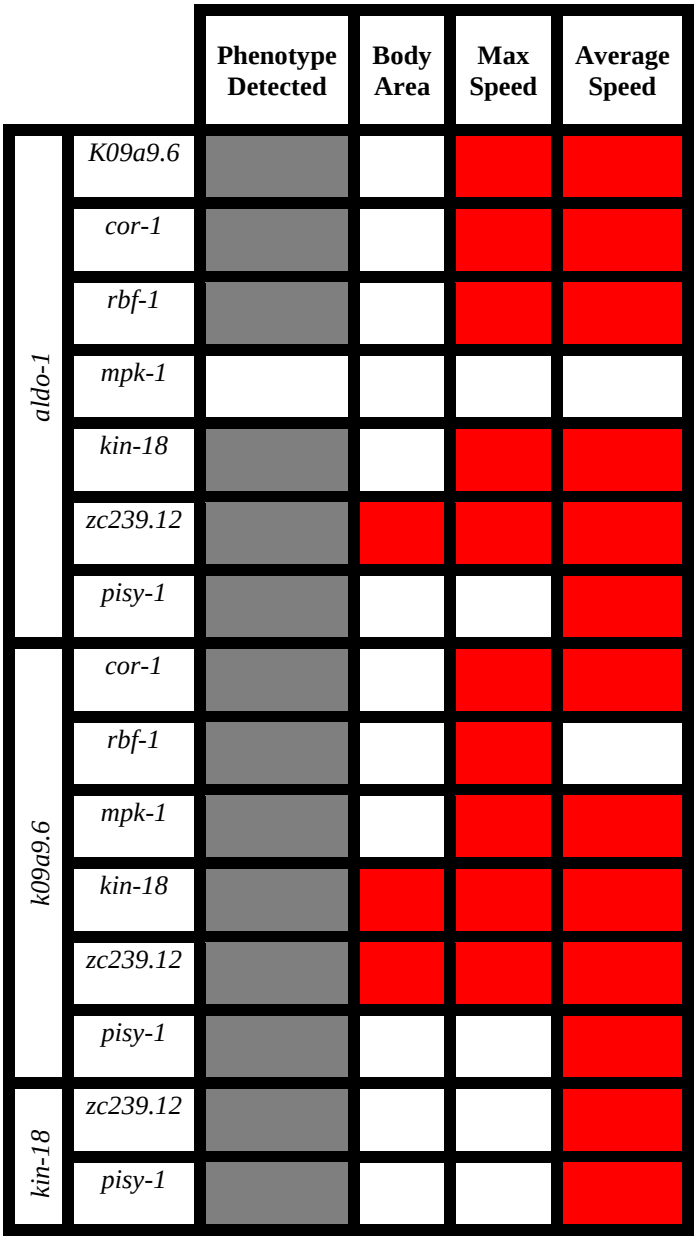
Two-Hit Gene Knockdown Results

After the series of one-hit knockdowns were performed and analyzed, each of the 8 genes were paired and mixed sequentially with each other to perform two-hit knockdown assays. Two-hit combinations can be seen in Table 3.

	aldo-1	***k09a9.6***	***cor-1***	***rbf-1***	***mpk-1***	***kin-18***	***zc239.12***	***pisy-1***
aldo-1		*aldo-1/k09a9.6*	*aldo-1/cor-1*	*aldo-1/rbf-1*	*aldo-1/mpk-1*	*aldo-1/kin-18*	*aldo-1/zc239.12*	*aldo-1/pisy-1*
k09a9.6			*k09a9.6/cor-1*	*k09a9.6/rbf-1*	*k09a9.6/mpk-1*	*k09a9.6/kin-18*	*k09a9.6/zc239.12*	*k09a9.6/pisy-1*
cor-1				*cor-1/rbf-1*	*cor-1/mpk-1*	*cor-1 /kin-18*	*cor-1/zc239.12*	*cor-1 /pisy-1*
rbf-1					*rbf-1/mpk-1*	*rbf-1/kin-18*	*rbf-1/zc239.12*	*rbf-1/pisy-1*
mpk-1						*mpk-1/kin-18*	*mpk-1/zc239.12*	*mpk-1/pisy-1*
kin-18							*kin-18/zc239.12*	*kin-18/pisy-1*
zc239.12								*zc239.12/pisy-1*
pisy-1								

Table 3: Two-hit Knockout Combinations of 16p11.2 genes

The two-hit model was used in addition to single-hit analysis to characterize the potential interplay of candidate orthologous genes within the human 16p11.2 chromosomal region. Three trials of each of the 28 respective two-hit knockdowns were performed and were characterized using the same assays performed with the one-hit populations. A fold change was calculated to normalize the values among each assay. The fold change calculation divided individual data points with the average wild-type values measured from each assay. Fold change analysis was performed with each two-hit assay. The first two trials of the two-hit knockout models were compared to the wild-type control, and these trends can be observed in Table 4. A red label was placed if significant values were observed in both trials of the respective two hit knockdown.



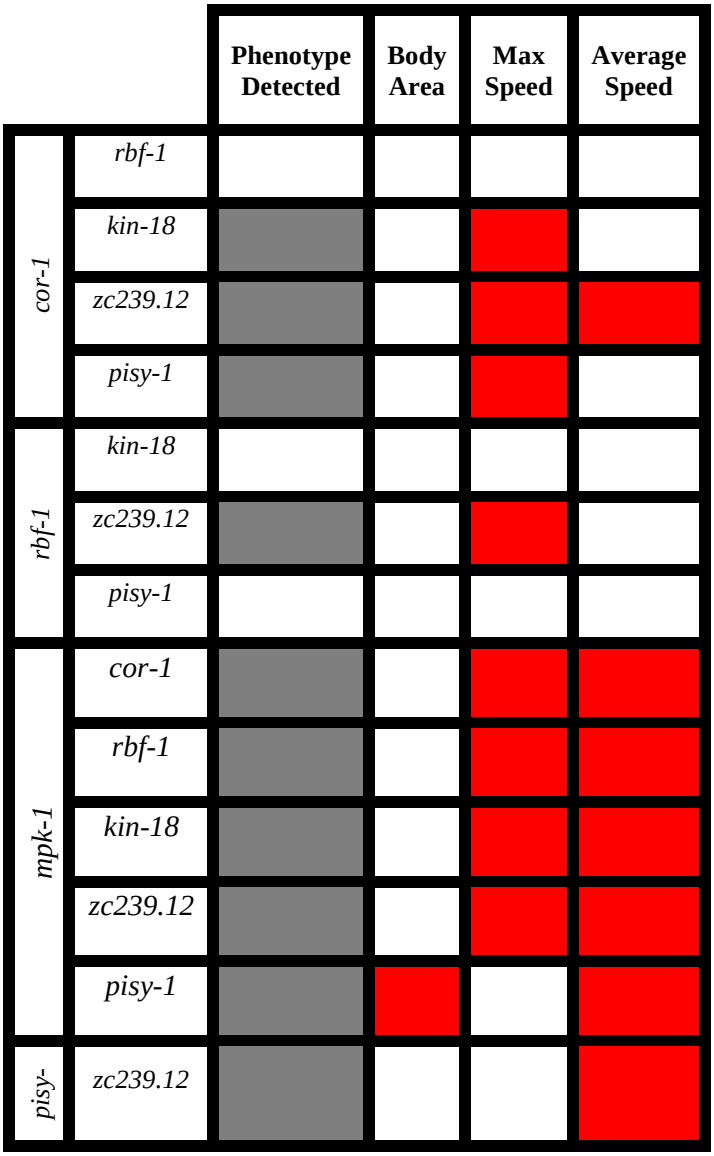


Table 4: Phenotypic trends associated with two-hit RNAi Treatment of 16p11.2 Orthologs from Trial 1 & 2 **Red Spaces indicate that two out of the two trials had data sets that possessed a p value of less than 0.05. Genes that are labeled as gray illustrate than an assay demonstrated a significant phenotype in at least two out of the three assays conducted.**

Body Area - Two Hits, Trial 1 & 2

The first two trials were conducted in order to observe for notable trends and potentially identify a relationship in reference to the wild-type (EV) control. Across the three assays performed, notable phenotypes can be primarily seen among two-hit knockouts made with *aldo-1*, *mpk-1*, and *k09a9.6* combinations. Upon seeding L1 worms onto each respective two-hit RNAi treated E. coli, C. elegans populations were imaged across the L1, L2, L3 and L4 growth stage in an identical procedure to that of the single-hit assay. Significance and trends were most noticeable in the L3 and L4 stage, and the fold change values for the L3 stage can be seen in Figure 6.

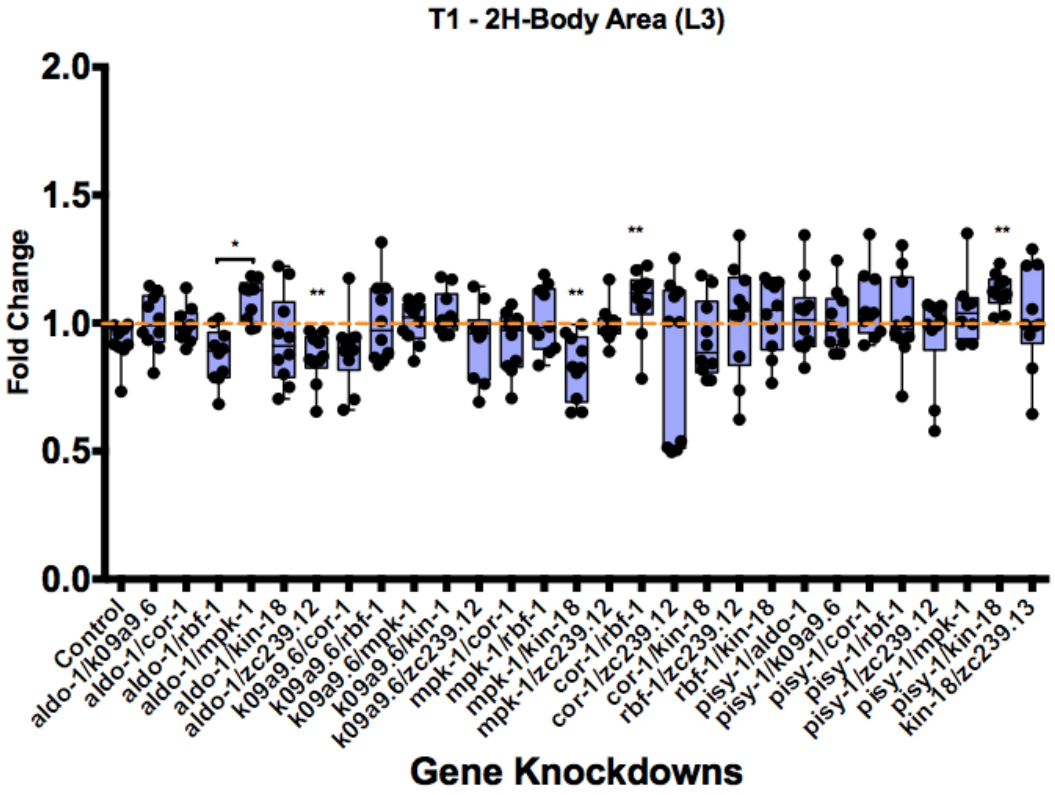


Figure 7: Trial 1 Two-Hit Knockdown Body Area (L3) (N = 10) **Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and **** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

In the L3 stage, *aldo-1/rbf-1*, aldo-1/zc239.12, and *mpk-1/kin-18* had body area measurements that were smaller than the wild-type. In contrast, two-hit knockouts of *aldo-1/mpk-1*, *cor-1/rbf-1*, and *pisy-1/kin-18* had body area measurements that were significantly larger than that of the wild-type. In comparison to the L3 stage, knockouts of the L4 stage enhanced the trends first observed in the former stage, and as a result, had more significant deviance from the wild-type vector. This pattern in the L4 growth stage can be visualized in Figure 7.

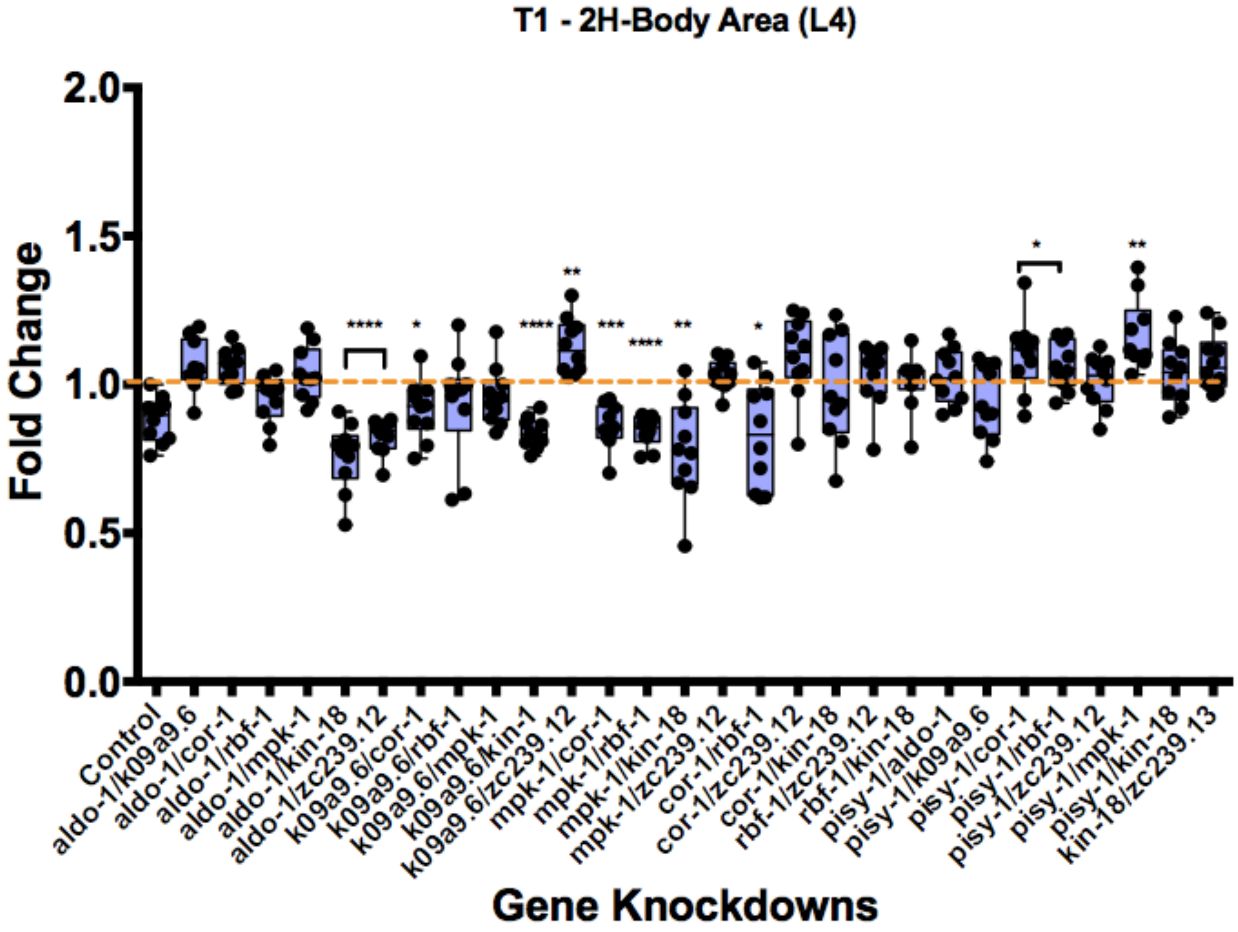


Figure 8: Trial 1 Two-Hit Knockdown Body Area (L3) (N = 10) **Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and **** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

Within the L4 stage, significant phenotypes can be best identified with two-hit knockouts paired with the knockout of *aldo-1, k09a9.6,* and *mpk-1*. In reference to the wild-type, it can be seen that *aldo-1/kin-18, aldo-1/zc239.12, k09a9.6/kin-18, mpk-1/cor-1, mpk-1/rbf-1, mpk-1/kin-18,* and *cor-1/rbf-1* were smaller than the wild-type. Smaller worms identified within these combinations contrast the trends observed from the single-hit assay, as knockouts of *aldo-1, k09a9.6, mpk-1,* and *kin-18* produced significantly larger worms. Two-hit knockdowns that contributed to larger worms in comparison to the wild-type were *k09a9.6/zc239.12, pisy-1/cor-1, pisy-1/rbf-1,* and *pisy-1/mpk-1*.

Crawling Speed - Two Hits, Trial 1 & 2

Two-hit knockout populations were recorded for 10 seconds and were measured for their average and max crawling speeds. Approximately 5-10 worms in the L4 stage were recorded in Trial 1, and the fold change of the values in comparison to the EV control can be observed in Figure 8. For the two-hit model, the crawling speed assays demonstrated the greatest significance in deviance from the wild-type.

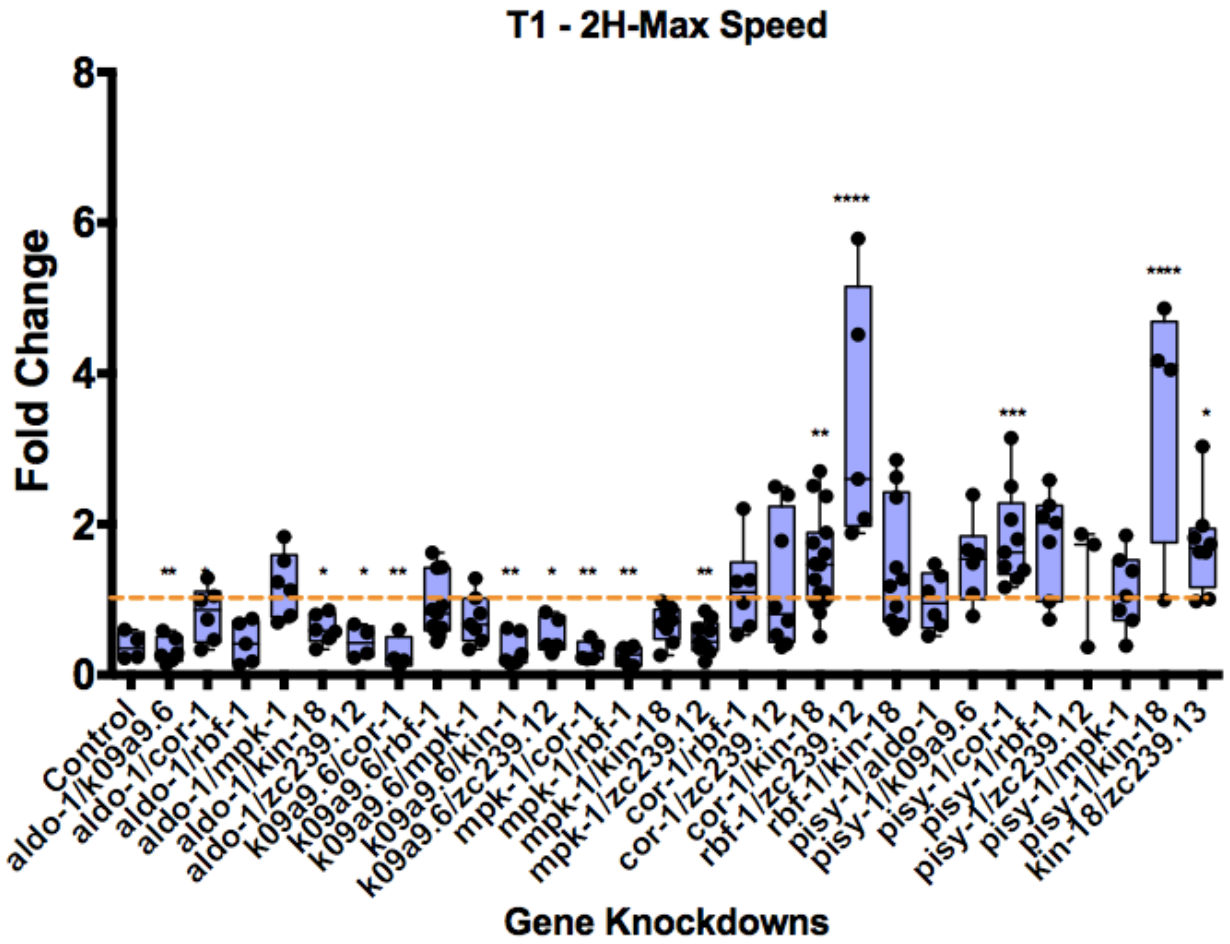


Figure 9: Trial 1 Two-Hit Knockdown Max Speed **Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and **** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

Combinations made with knockouts of the *aldo-1, k09a9.6,* and *mpk-1* gene were characterized as having slow max speed values. These trends are consistent with the trials performed during the one-hit trials, as populations affected by the *aldo-1, k09a9.6,* and *mpk-1* knockout were significantly slower than the knockout population. However, it is notable that two hit knockouts of *aldo-1/mpk-1* & *k09a9.6/rbf-1* had values similar to the wild-type measurements. Interestingly, two-hit knockdowns involving *pisy-1* produced individuals that had the greatest max speeds.

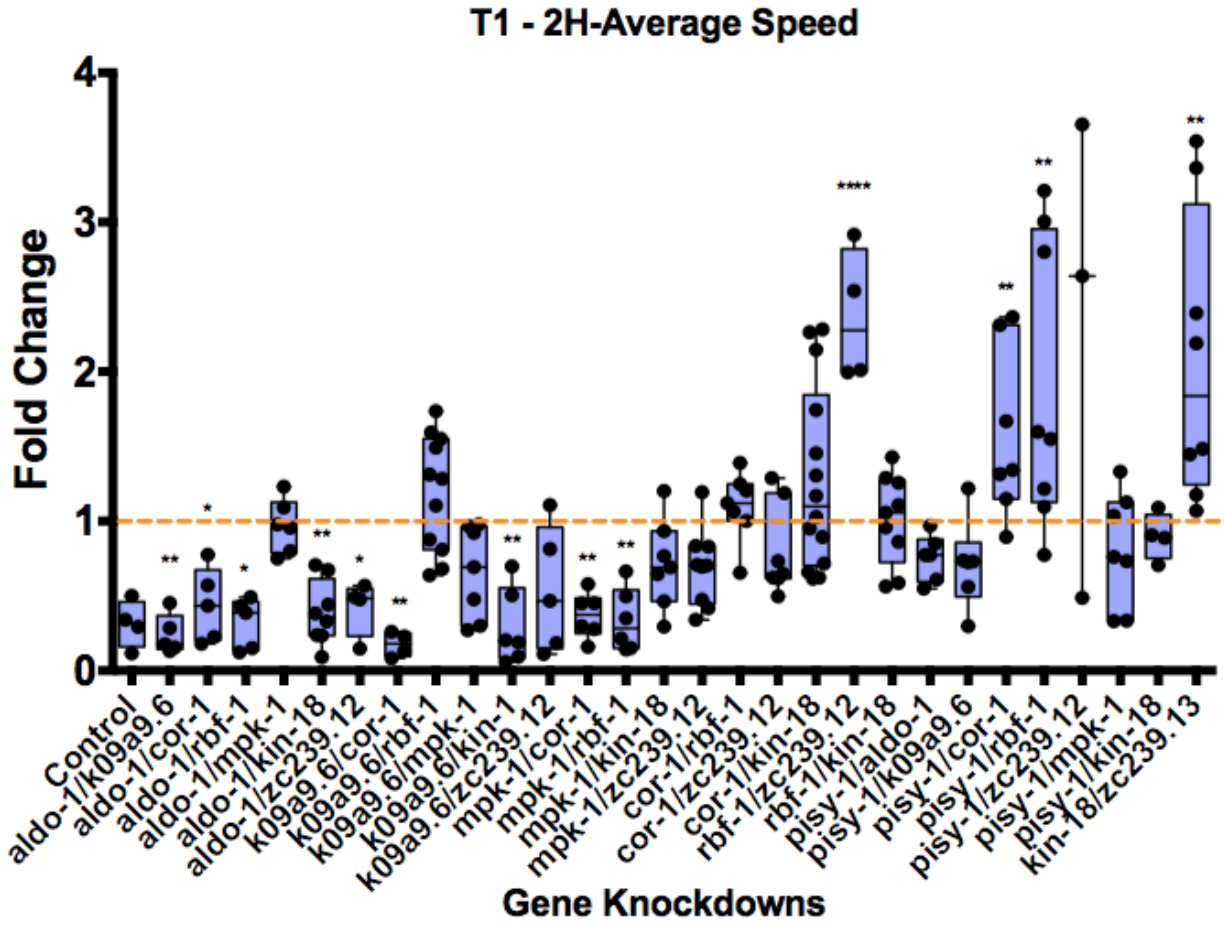


Figure 10: Trial 1 Two-Hit Knockdown Average Speed Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and ** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

Average speed values mirror the trends observed within Figure 8. Combinations made with knockouts of the *aldo-1*, *k09a9.6*, and *mpk-1* gene were again characterized as having slower speeds, while knockouts crossed with *pisy-1* greatly increased average crawling speeds.

Body Area - Two Hits, Trial 3

The first two trials performed with the two-hit populations observed for the deviance of phenotypes in reference to the wild-type. However, the notable difference between the two-hit model vs the one-hit model is the RNAi dosage in effect to gene knockdown. Within the third trial, to maintain the 1:1 distribution of knockdown through RNAi feeding, half the volume of each RNAi-treated E. coli was given to reach the volume used during Assay 1 and the single hit assays. As a result, half the concentration of each single-hit RNAi gene is present within the two-hit models. After confirming the validity and effects of the two-hit assay, trial 3 was conducted with the two-hit combinations and two respective negative controls: EV + First Knockout and EV + Second Knockout. This allowed a justifiable comparison to be made between the two-hit model in comparison to the trends seen from the one-hit knockdowns, as the concentrations of each component would be consistent.

In Trial 3, combinations were performed in batches and fold change values were calculated in reference to both controls (EV/First Knockout Gene) and (EV/Second Knockout Gene). Body area was measured across the L1-L4 growth stage, and fold change analysis is depicted for the L3 stage in Figure 10.

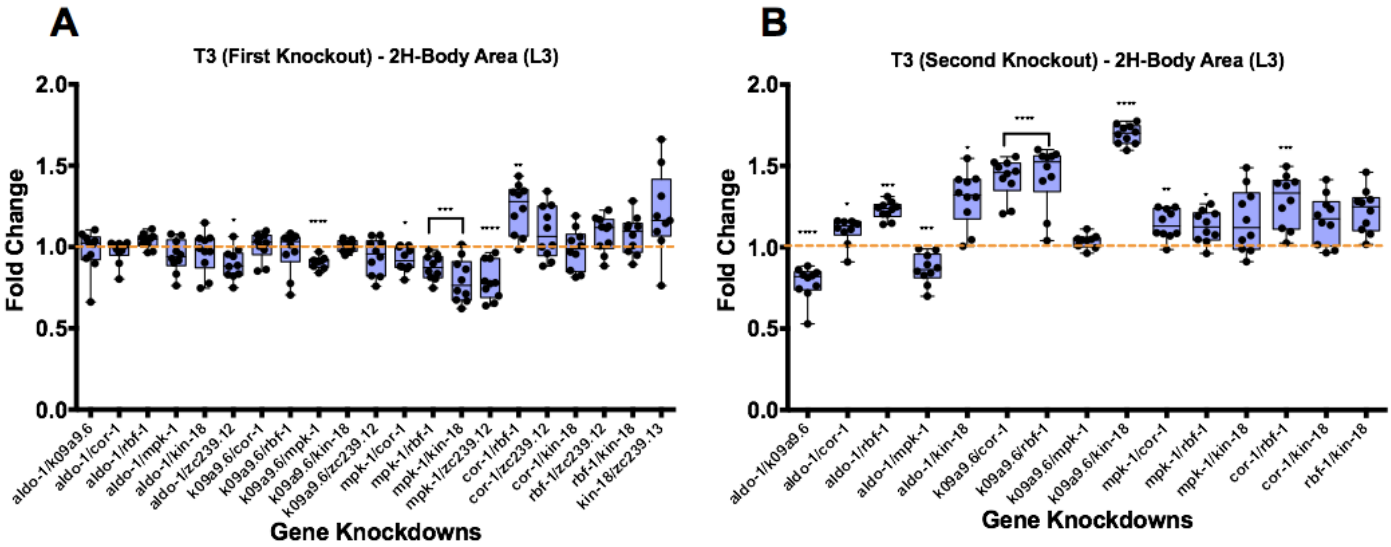


Figure 11: Trial 3 Two-Hit Knockdown Body Area (L3) (N = 10)A. Represents fold change values of 2-Hit knockdowns in comparison to EV/*First gene knockdown.* B. Represents fold change values of 2-Hit knockdowns in comparison to EV/*Second gene knockdown.* Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and ** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

In the L3 phase, *aldo-1/k09a9.6* knockouts were smaller than EV/*aldo-1*. In comparison to EV/*k09a9.6*, two-hit knockouts of *k09a9.6/mpk-1* and *aldo-1/k09a9.6* were significantly smaller. Two-hit genes containing a *mpk-1* knockout were generally smaller than the EV/*mpk-1* control. In contrast, two-hit genes containing a *cor-1* knockout were generally larger than the EV/*cor-1* control. Two-hit knockout combinations conducted with a *rbf-1* knockout produced increased body area in *aldo-1/rbf-1, k09a9.6/rbf-1,* and *mpk-1/rbf-1* in comparison to the EV/*rbf-1* control. Two-hit knockout with *kin-18* did not produce any significant trends when compared to the EV/*kin-18* knockout.

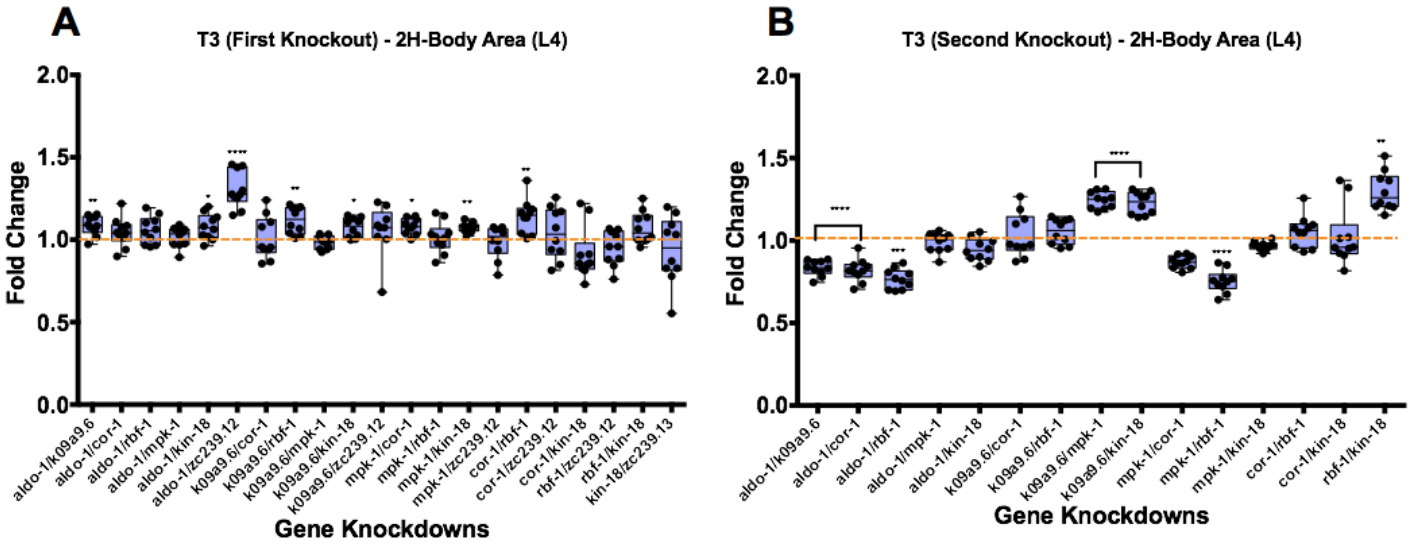


Figure 12: Trial 3 Two-Hit Knockdown Body Area (L3) (N = 10) **A. Represents fold change values of 2-Hit knockdowns in comparison to EV/*First gene knockdown.* B. Represents fold change values of 2-Hit knockdowns in comparison to EV/*Second gene knockdown.* Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and **** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

In the L4 phase, there was greater deviance across all two-hit knockouts to their respective controls. In comparison to EV/*aldo-1*, *aldo-1/k09a9.6*, *aldo-1/kin-18*, and *aldo-1/zc239.12* were significantly larger in body area. Two-hit populations were variable in reference to the EV/*k09a9.6* control, as knockouts of *k09a9.6/rbf-1* and *k09a9.6/kin-18* demonstrated larger worms, while *aldo-1/k09a9.6* had smaller measured body areas. Two-hit genes containing a *mpk-1* knockout were generally larger than the EV/*mpk-1* control, as significance was noted with *mpk-1/cor-1*, *mpk-1/kin-18*, and *k09a9.6/mpk-1*. Two-hit genes containing a *cor-1* knockout were generally smaller in comparison to the EV/*cor-1* control, although the *cor-1/rbf-1* had larger body area measurements. In reference to EV/*rbf-1*, *aldo-1/rbf-1* and *mpk-1/rbf-1* were smaller in body area than the control. Two-hit knockout with *kin-18* were generally larger than the *EV/kin-18* control, as indicated by the two-hit knockouts of *k09a9.6/kin-18* and *rbf-1/kin-18*.

Crawling Speed - Two Hits, Trial 3

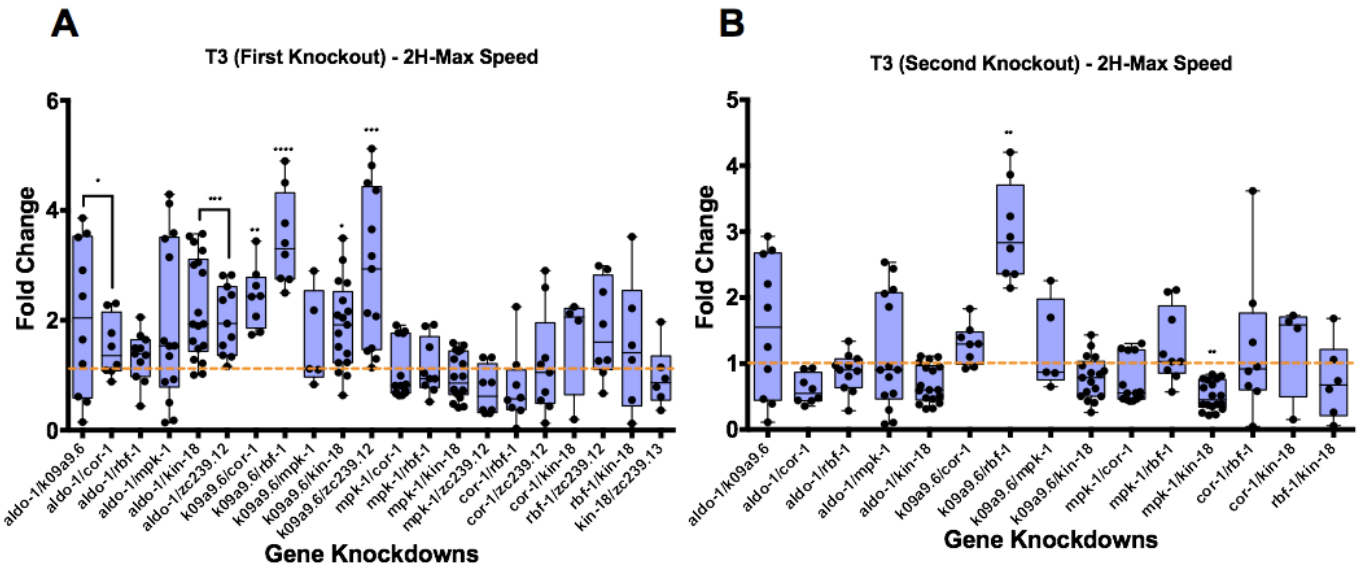


Figure 13: Trial 3 Two-Hit Knockdown Average Speed **A. Represents fold change values of 2-Hit knockdowns in comparison to EV/*First gene knockdown.* B. Represents fold change values of 2-Hit knockdowns in comparison to EV/*Second gene knockdown.* Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and **** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

Analysis of the max speed for the two-hit knockdowns provided some unique trends. In comparison to the EV/*aldo-1* and EV/*k09a9.6* controls, knockouts containing *aldo-1* and *k09a9.6* were faster in max speeds. Two-hit knockouts with *mpk-1* did not produce any significant trends when compared to the EV/*mpk-1* knockout. Two-hit knockouts of k09a9.6/rbf-1 and mpk-1/kin-18 were in faster and slower in relation to their respective control. However, two-hit knockouts containing *cor-1*, *rbf-1*, and *kin-18* in general did not demonstrate any significance in relation to EV/*cor-1*, EV/*rbf-1*, and EV/*kin-18*.

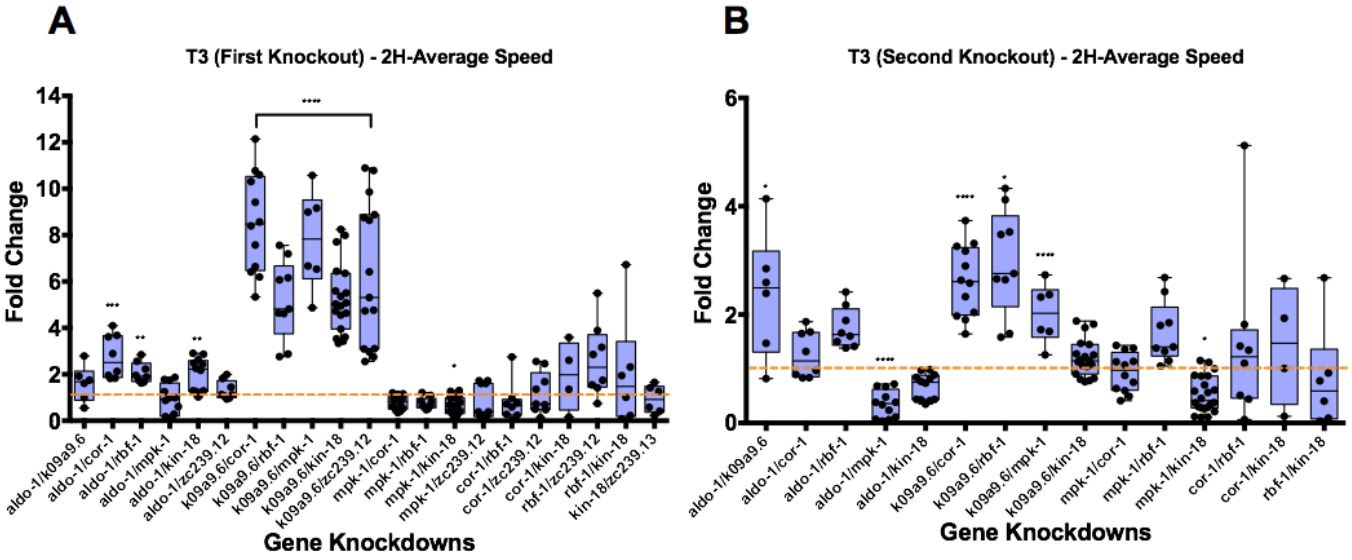


Figure 14: Trial 3 Two-Hit Knockdown Average Speed **A. Represents fold change values of 2-Hit knockdowns in comparison to EV/*First gene knockdown.* B. Represents fold change values of 2-Hit knockdowns in comparison to EV/*Second gene knockdown.* Boxes indicate the upper and lower quartiles, and the line indicates the median. Error bars indicate the maximum and minimum within the population. *, **, ***and **** represent p <0.05, p<0.01, p<0.001, and p<0.0001. Significance was calculated using Student T-test.**

Similar to the analysis of the max speed for the two-hit knockdowns, measurement of the average speed indicated consistent unique trends. Populations containing two-hit knockouts of *aldo-1* and *k09a9.6* were in general faster than the EV/*aldo-1* and EV/*k09a9.6* controls. The two-hit knockouts containing *k09a9.6* were the most notable however, as all the two-hit knockouts featuring an *k09a9.6* gene demonstrated an increase of average speed in comparison to the EV/*k09a9.6* control. As a result, populations with a two-hit knockout that included the *k09a9.6* cross were seen to have a rescue phenotype, as single-hit analysis of *k09a9.6* knockouts produced significantly slower worms. 2-hit knockouts of *mpk-1/kin-18* and *aldo-1/mpk-1* were significantly slower than the EV/*mpk-1* control, while *k09a9.6/mpk-1* was seen to have a greater degree of average speed. 2-hit knockouts containing *cor-1*, *rbf-1*, and *kin-18* were not seen to be significant

unless when knocked down with *k09a9.6* and *mpk-1* in relation to their respective EV/*knockout* controls.

Chapter 4 : Discussions and Conclusions

In this study, I analyzed the effects of one-hit knockdowns within orthologous genes of the 16p11.2 region on body area and neurodevelopment. Eight genes were knocked down through RNAi induced silencing, and phenotype variations were characterized using a fold change of individual data points with respect to the wild-type. Of the eight genes characterized, when *aldo-1, k09a9.6* and *mpk-1* were knocked down, more pronounced phenotypes (i.e. larger worms and slower crawling speeds) were observed across three trials. My data emphasizes the importance of the genes orthologous to the 16p11.2 region as modulators of neurodevelopment and motor function. This study has the potential to broaden our ability to target these genes for new therapies for individuals diagnosed with autism spectrum disorders (ASDs).

To discover the function of *aldo-1, k09a9.6,* and *mpk-1*, I investigated how knockdowns of these genes, along with the five other 16p11.2 orthologs mentioned in Table 2. Body area growth and crawling speeds were measured across three trials. Across the L1-L4 growth period, *C. elegans* affected with the *aldo-1, k09a9.6, kin-18,* and *mpk-1* knockout were characteristic of being smaller than the wild-type vector prior to the final L4 stage. This finding is often seen among individuals with ASDs, as Autism phenotypes are often defined by development delay and stunted growth.[24]

[24] Tanaka AJ et al. De novo variants in EBF3 are associated with hypotonia, developmental delay, intellectual disability, and autism. Cold Spring Harb Mol Case Stud. 2017 Nov 21;3(6). pii: a002097. doi: 10.1101/mcs.a002097. Print 2017 Nov. PubMed PMID: 29162653; PubMed Central PMCID: PMC5701309.

However, by the L4 stage, the body area of the worms had a significant increase in comparison to the wild-type strain. It has been noted that human subjects heterozygous for deletions at the 16p11.2 region displayed high penetrance for obesity.[25] The increase in body area of these knock out populations could be related to the suppression of the normal functionality of the *aldo-1, mpk-1,* and *kin-18* genes. The *aldo-1, mpk-1,* and *kin-18* genes mediate elongation of the worm embryo, cellular proliferation, and feeding behavior respectively. Suppression of these genes could be responsible for the aberrant phenotypes observed at the L4 stage. Although having similar phenotypic trends, knockouts of the remaining four genes did not differ in body area to a significant degree. Curiously, preliminary data shows that one-hit models had a significant decrease in average and max speed in relation to the wild-type strain across the three trials. Knockouts of the *aldo-1, k09a9.6, cor-1, rbf-1,* and *mpk-1* genes were significant across each of the three trials. The *cor-1, rbf-1,* and *kin-18* genes are involved in cytoskeleton development, motor neuron development, and nematode locomotion respectively. The *k09a9.6* and *zc239.12* genes are uncharacterized in the *C. elegans* model function. These findings could suggest that a knockout of these genes within the 16p11.2 region affect the development and neuromotor pathway. If this is the case, it would be worth investigating the knockdown and effect of two orthologous genes to gain an understanding of phenotypic variations.

The two-hit knockdown model in *C. elegans* populations paired the eight available orthologous genes and measured body area across the L1-L4 growth stages, and crawling speed in

[25] Walters RG et al. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. Nature. 2010 Feb 4;463(7281):671-5. doi: 10.1038/nature08727. PubMed PMID: 20130649; PubMed Central PMCID: PMC2880448.

the L4 growth stage. The first two trials showed that two-hit knockdowns with *aldo-1/kin-18, aldo-1/zc239.12, k09a9.6/cor-1, k09a9.6/kin-18, mpk-1/cor-1, mpk-1/kin-18* and *mpk-1/rbf-1* had phenotypes associated with smaller body area. This was interesting because within the single-hit assay, *aldo-1, k09a9.6,* and *mpk-1* were associated with larger body area. Measurements of crawling speeds were consistent to what was seen during the one hit assays for two-hits containing pairings with *aldo-1, k09a9.6,* and *mpk-1*. These two-hit knockout populations were generally seen to be slower in both max speed and average speed. However, two-hit knockouts containing *rbf-1* and *pisy-1* were seen to have worms that were faster than the wild-type population.

The third trial in the 2-hit knockdown study was performed to measure trends in reference to one hit knockouts, as opposed to measuring the trends in relation to the wild-type population. In the L3 phase, it was seen that in reference to EV/*aldo-1,* only the *aldo-1/k09a9.6 2*-hit knockout were smaller than the control. In comparison to EV/*k09a9.6,* two-hit knockouts of *k09a9.6/mpk-1* and *aldo-1/k09a9.6* were significantly smaller. In general, it was seen that two-hit knockouts containing a *mpk-1* knockout were generally smaller than the EV/*mpk-1* control. These results correlate with what was originally hypothesized, in that two-hit knockdowns cause an increase in the general trend first seen in the single-hits. In contrast, two-hit genes containing a *cor-1, rbf-1,* and a *kin-18* knockout were generally larger than the respective EV/knockout control. This was interesting, since in the single-hit assay, *cor-1* and *rbf-1* knockouts did not display any significant trends in body area sizes in the L3 stage. In the L4 phase, there was greater deviance across all two-hit knockouts to their respective controls. In comparison to EV/*aldo-1,* two-hit knockouts containing *aldo-1* were larger in body area. Two-hit populations were variable in reference to the EV/*k09a9.6* control, as knockouts of *k09a9.6/rbf-1* and *k09a9.6/kin-18* demonstrated larger

worms. However, the two-hit knockout of *aldo-1/k09a9.6* displayed a rescue phenotype in which the population had significantly smaller body area measurements in reference to the EV/*k09a9.6* control. Two-hit genes containing a *mpk-1* knockout were consistent with one-hit trials, as knockout populations were generally larger in body size than the EV/*mpk-1* control.

Analysis of the max speed and average speed for the two-hit knockouts in trial 3 in general did not provide the expected trends predicted from trail 1. 2-hit knockouts containing *aldo-1* and *k09a9.6* were significantly faster in crawling speeds in comparison to the EV/*aldo-1* and EV/*k09a9.6* controls. This is notable, as it contrasts the one-hit trials, in which single-hits of *aldo-1* and *k09a9.6* caused a significant decrease in the crawling speeds of those populations. This rescue phenotype was not seen with two-hit knockouts containing the *mpk-1* gene in reference to the EV/*mpk-1* knockout. Two-hit combinations with the other genes did not display significant variance from their respective EV/knockout controls.

Overall this study has shown that knockouts of orthologous genes within the 16p11.2 region provide significant phenotypic trends in body area and crawling speed in relation to the wild-type control. Two-hit knockouts provided expected trends when referenced to the wild-type control, as populations crossed with the *aldo-1, k09a9.6,* and *mpk-1* knockout generally displayed similar or more significant trends as seen with the one-hit trials. However, the direct comparison made between the two-hit knockouts and their respective single-hit controls elucidated the potential interplay associated with the complex network of neurons in the *C. elegans* model to deliver rescue phenotypes comparative to single-hit knockouts. Overall, the results of this study

suggest the continued importance of understanding the mechanistic factors associated with morphological changes and motor dysfunction within the 16p11.2 region.

Chapter 5 Bibliography

1. Bijlsma EK et al. (2009) Extending the phenotype of recurrent rearrangements of 16p11.2: deletions in mentally retarded patients without autism and in normal individuals. Eur J Med Genet 52:77-87.
2. Rosenfeld JA, Coppinger, J., Bejjani, B.A., Girirajan, S., Eichler, E.E., Shaffer, L.G., Ballif., B.C. (2010) Speech delays and behavioral problems are the predominant features in individuals with developmental delays and 16p11.2 microdeletions and microduplications. J Neurodevelop Disord 2:26-38.
3. Shinawi M, Liu P, Kang SH, Shen J, Belmont JW, Scott DA, et al. Recurrent reciprocal 16p11.2 rearrangements associated with global developmental delay, behavioural problems, dysmorphism, epilepsy, and abnormal head size. Journal of medical genetics. 2010;47(5):332–41. Epub 2009/11/17.
4. Blomquist HK, et al. Frequency of the fragile X syndrome in infantile autism. A Swedish multicenter study. Clin. Genet. 1985;27:113–117.
5. Sebat J, et al. Strong association of *de novo* copy number mutations with autism. Science. 2007;316:445–449. Beyond identifying important CNV likely to prove important to our understanding of the ASDs, this study highlights a significant difference in the frequency of *de novo* variants between simplex and multiplex families, raising the possibility that distinct mechanisms are involved in each.
6. Jacquemont ML, et al. Array-based comparative genomic hybridisation identifies high

frequency of cryptic chromosomal rearrangements in patients with syndromic autism spectrum disorders. J. Med. Genet. 2006;43:843–849. The numerous *de novo* deletions that are reported in this work have received relatively little attention but are probably important in the ASDs. These data also show that rare *de novo* mutations are likely to appear at particularly high frequencies in syndromic populations.

7. Abrahams BS, Geschwind DH. Advances in autism genetics: on the threshold of a new neurobiology. Nat Rev Genet. 2008;9:341–55. doi: 10.1038/nrg2346.
8. Kumar RA, KaraMohamed S, Sudi J, Conrad DF, Brune C, et al. Recurrent 16p11.2 microdeletions in autism. Hum Mol Genet. 2008;17:628–38. doi: 10.1093/hmg/ddm376.
9. Rosenfeld, J. A., Coppinger, J., Bejjani, B. A., Girirajan, S., Eichler, E. E., Shaffer, L. G., & Ballif, B. C. (2010). Speech delays and behavioral problems are the predominant features in individuals with developmental delays and 16p11.2 microdeletions and microduplications. *Journal of Neurodevelopmental Disorders*, *2*(1), 26–38.
10. Mefford HC et al. (2010) Genome-wide copy number variation in epilepsy: novel susceptibility loci in idiopathic generalized and focal epilepsies. PLoS Genet 6:e1000962.
11. Inoue T, Yatsuki H, Kusakabe T, Joh K, Takasaki Y, Nikoh N, Miyata T, Hori K. Caenorhabditis elegans has two isozymic forms, CE-1 and CE-2, of fructose-1,6-bisphosphate aldolase which are encoded by different genes. Arch Biochem Biophys. 1997 Mar 1;339(1):226-34. PubMed PMID: 9056253.

12. Wang X, Zhou F, Lv S, Yi P, Zhu Z, Yang Y, Feng G, Li W, Ou G. Transmembrane protein MIG-13 links the Wnt signaling and Hox genes to the cell polarity in neuronal migration. Proc Natl Acad Sci U S A. 2013 Jul 2;110(27):11175-80. doi: 10.1073/pnas.1301849110. Epub 2013 Jun 19. PubMed PMID: 23784779; PubMed Central PMCID: PMC3703986.

13. Staunton J, Ganetzky B, Nonet ML. Rabphilin potentiates soluble N-ethylmaleimide sensitive factor attachment protein receptor function independently of rab3. J Neurosci. 2001 Dec 1;21(23):9255-64. PubMed PMID: 11717359.

14. Schouest KR, Kurasawa Y, Furuta T, Hisamoto N, Matsumoto K, Schumacher JM. The germinal center kinase GCK-1 is a negative regulator of MAP kinase activation and apoptosis in the C. elegans germline. PLoS One. 2009 Oct 14;4(10):e7450. doi: 10.1371/journal.pone.0007450. PubMed PMID: 19826475; PubMed Central PMCID: PMC2757678.

15. Berman KS, Hutchison M, Avery L, Cobb MH. kin-18, a C. elegans protein kinase involved in feeding. Gene. 2001 Nov 28;279(2):137-47. PubMed PMID: 11733138; PubMed Central PMCID: PMC4441751.

16. Croll N (1975) Components and patterns in the behavior of the nematode *Caenorhabditis elegans*. J Zool 176: 159–176.

17. Park SM, Littleton JT, Park HR, Lee JH. Drosophila Homolog of Human KIF22 at the Autism-Linked 16p11.2

18. Loci Influences Synaptic Connectivity at Larval Neuromuscular Junctions. Exp Neurobiol. 2016 Feb;25(1):33-39.

19. Jasmine M. McCammon, Alicia Blaker-Lee, Xiao Chen, Hazel Sive; The 16p11.2

homologs fam57ba and rbf-1 generate certain brain and body phenotypes, Human Molecular Genetics, Volume 26, Issue 19, 1 October 2017, Pages 3699–3712

20. White, J.G., E. Southgate, J.N. Thompson, and S. Brenner. 1986. The structure of the nervous system of the nematode Caenorhabditis elegans. Philos. Trans. R. Soc. Lond. B Biol. Sci. 314: 1-340.

21. Durbin, R.M. 1987. Studies on the Development and Organisation of the Nervous System of Caenorhabditis elegans. In. Cambride University, Cambridge.

22. Croll, N.A. 1975. Components and patterns in the behaviour of the nematode Caenorhabditis elegans. J. Zool. Lond 176: 159-176.

23. Calixto A, Chelur D, Topalidou I, Chen X, Chalfie M. Enhanced neuronal RNAi in C. elegans using SID-1. Nature methods. 2010;7(7):554–9. pmid:20512143

24. Chen, P., Wang, D., Chen, H., Zhou, Z., & He, X. (2016). The nonessentiality of essential genes in yeast provides therapeutic insights into a human disease. *Genome Research*, *26*(10), 1355–1362. http://doi.org/10.1101/gr.205955.116

25. Tanaka AJ et al. De novo variants in EBF3 are associated with hypotonia, developmental delay, intellectual disability, and autism. Cold Spring Harb Mol Case Stud. 2017 Nov 21;3(6). pii: a002097. doi: 10.1101/mcs.a002097. Print 2017 Nov. PubMed PMID: 29162653; PubMed Central PMCID: PMC5701309.

26. Walters RG et al. A new highly penetrant form of obesity due to deletions on chromosome 16p11.2. Nature. 2010 Feb 4;463(7281):671-5. doi: 10.1038/nature08727. PubMed PMID: 20130649; PubMed Central PMCID: PMC2880448.

Academic Vita of Ayush Thomas

amt5650@psu.edu

EDUCATION

Schreyer Honors College, The Pennsylvania State University Class of 2018

B.S., Eberly College of Science
Major: Biochemistry and Molecular Biology
Minor: Microbiology, Spanish
Honors in Biochemistry and Molecular Biology

PROFESSIONAL EXPERIENCE

The Girirajan Lab, Department of Biochemistry and Molecular Biology
Undergraduate Researcher

- Developed functional genomic research models and strategies to observe gene function within high-risk chromosome regions associated with neurodevelopmental disorders, including, Autism and Schizophrenia
- Conducted two independent genomic projects, gaining technical skills in protein purification, enzyme specific activity assays, in vivo experimentation, and fluorescence microscopy
- Modeled individual biochemistry associated mechanistic pathways associated with loss of gene function in C. elegans

Achaogen
Biosynthesis Research Intern in Antibacterial Drug Development

- Designed and optimized a procedure to adapt known antibiotic pathways to generate novel second-generation antibiotics immune to growing bacterial resistance mechanisms
- Consulted with various protein chemists and organic chemists to evaluate technologies to accelerate internal research efforts
- Presented project results to the early development research team and company stakeholders at a web conference

Penn State Learning
Organic Chemistry Guided Study Group Leader

- Teach two weekly 90-minute study sessions available to over 800 students, attended by up to 200 students per session
- Prepare PowerPoint presentations with practice problems and strategies for students

LEADERSHIP EXPERIENCE

Science Lion Pride
Alumni Relations Director

- Conduct alumni and donor outreach through collaboration with the Eberly College of Science alumni

- Represent the Eberly College of Science during Alumni events to discuss the progress Eberly's academic programs and research advancement

Scholar Advancement Team
Co-Director

- Represent over 1,200 Schreyer Scholars to further the vision of The Honors College amongst prospective students, alumni visitors, and current and selected donors
- Provide a student perspective at various events such as the Schreyer Alumni Reunion, the Outstanding Scholar Alumni Award Ceremony, Faculty Awards Ceremony lectures, Distinguished Speaker Series lectures, and other functions

Meet the Professor Club
President

- Collaborate with professors from the Eberly College of Science to host bi-weekly lunch discussions with undergraduate students
- To date, 20 professors and over 75 students have participated

TEDxPSU
Director of Partnership

- Led a partnerships team of 5 students that communicated with 10 corporate partners, and raised over $55,000 dollars for the conference.
- Marketed brand and created partnerships with local businesses, student organizations, and colleges

---

HONORS

The Pennsylvania State University Academic Excellence Scholarship
Charles and Vickie Grier Undergraduate Research Award