Functional Assessment of Human 21st Chromosome Orthologs in *Caenorhabditis elegans*

Presented by Sofia Smith

In partial fulfillment of the requirements for graduation with the Dean's Scholars Honors Degree in Biology

_____	_____
Jonathan Pierce, Ph.D. Department of Neuroscience Supervising Professor	Date
_____	_____
Shelley Payne, Ph.D. Honors Advisor in Biology	Date

Abstract

Down syndrome (DS) is a neurodevelopmental disorder caused by having an extra copy of the 21st chromosome. It is the most common genetic cause of intellectual disability (ID), occurring in 1 in 700 live births. DS involves a variety of phenotypes, but every person with DS exhibits muscle weakness and ID. The extra copy of chromosome 21 includes nearly 200 protein-coding genes, and causation can be challenging to pinpoint as many elements of cellular biology are altered. DS researchers aim to elucidate the genotype to phenotype relationship of the syndrome, to explain why having an extra 21st chromosome causes the syndromic effects that it does.

Many studies of DS have investigated regions of the 21st chromosome at a time, grouping genes together by location. However, little is known about the individual genes that make up the human 21st chromosome. Establishing which genes contribute to DS could clarify the mechanism for the neuromuscular symptoms associated with DS. In attempting to fill this gap, I have worked with the laboratory of Dr. Jon Pierce to systematically characterize the genes of the 21st chromosome for a potential role in the nervous system or muscle function.

To this end, the Pierce lab uses the nematode *Caenorhabditis elegans* to study orthologs, or equivalent genes, from the human 21st chromosome. The general approach is to identify genes conserved between human and *C. elegans* and conduct broad, unbiased screens to identify potentially important genes. *C. elegans* can be a useful model for taking this approach to DS, as it has been used to discover the function of hundreds of genes in the nervous system and muscle in humans and shares orthologs with many of the protein-coding genes on our 21st chromosome.

We first conducted a loss-of-function screen, taking the *C. elegans* genes and systematically "breaking" them to observe phenotypic defects. The *C. elegans* loss-of-function mutants underwent a series of behavioral and pharmacological assays to evaluate phenotypic defects. The loss-of-function screen resulted in the identification of 10 genes thought to be critical for synaptic or motor function. Several of these genes were not previously identified for a contribution to DS.

We have since begun a gain-of-function screen using similar behavioral tests to empirically determine which genes are problematic when overexpressed. Preliminary results implicate a few genes in disrupting neuron-to-muscle circuitry. I have also verified transcriptional overexpression in mutants relative to control strains.

I aim to provide a more complete understanding of the individual genes involved in DS-associated phenotypes. Doing so can provide a foundation for other DS researchers and inform efforts toward a more precise mechanism of action. The hope is that these genes might later serve as therapeutic targets for alleviating neuromuscular symptoms in people with DS.

Introduction

Down syndrome (DS) is caused by a complete or partial trisomy of the 21st chromosome. DS is the most frequent survivable autosomal aneuploidy and leads to lifelong cognitive impairment. Each person with DS experiences phenotypes of intellectual disability (ID), such as learning and memory challenges, and muscle weakness which can make it difficult to speak and move efficiently. Other phenotypes with incomplete penetrance include heart defects, childhood leukemia, and early-onset Alzheimer's disease. Interestingly, people with DS have a lower risk of developing solid tumors (Antonarakis et al., 2020).

It was discovered over 60 years ago that DS is caused by an extra copy of the 21st chromosome (Human somatic autosome 21, HSA21), but it is unclear why the extra genetic material leads to DS. Each theory discussed here has merits and shortcomings, and it is likely that the connection between trisomy 21 and DS contains elements of both. One theory suggests that DS symptoms are caused by the cellular strain that results from extra genetic material (Patterson, 2009). It is true that there are direct and indirect disruptions on the rest of the genome and on cellular and tissue function. For example, trisomy of HSA21 causes global DNA hypermethylation on other chromosomes, which can affect the transcription of numerous genes (Antonarakis et al., 2020). However, comparing partial trisomies has shown that the degree of symptoms is not explained by the amount of extra genetic material (Lyle et al., 2009).

Another theory held that a "Down Syndrome Critical Region" (DSCR) that contains a select group of genes is responsible, rather than the burden of the chromosome itself. The DSCR theory was formed after mapping the area of chromosomal overlap between two partial trisomic patients that had DS, which resulted in a region of about 30 genes. Some evidence supports this theory as well. The Ts1Rhr mouse model, only trisomic for genes representing the human DSCR, displayed similar phenotypes to the Ts65Dn model that is trisomic for larger chromosomal regions (Belichenko et al., 2009). However, another study found that the Ts1Rhr trisomic DSCR genes are necessary, but not sufficient, to achieve the structural changes seen in Ts65Dn mice brains (Olson et al., 2007). Additionally, the partial trisomies of 30 people with DS were aligned, and there was not a region of overlap that invariably led to DS (Lyle et al., 2009). Despite agreement that the DSCR is likely to contain a few critical genes, the field lacks a broad consensus about the theory because the DSCR alone does not explain all symptoms. A recent study in favor of a more narrow region, the Highly Restrictive DSCR, renewed conversations about this theory (Pelleri et al., 2019).

Much of the foundational DS research was conducted in mouse models. Many useful models exist, such as the broadly used Ts65Dn mouse model of DS and a recent humanized mouse model that contains a nearly complete long arm of HSA21 (Kazuki et al., 2019). However, most of the work done in mice has looked at regions of the chromosome or sets of genes at a time. It can be beneficial to investigate individual genes through loss-of-function or gain-of-function analyses to dissect the genetic basis for a particular phenotype. It is also much simpler to design targeted therapeutics when there are individual candidate genes rather than large chromosomal regions. *C. elegans* is a useful model for conducting genetic screens, as it is simple to manipulate individual genes and conduct matings to achieve genetic crosses (Brenner, 1974).

The genetic screens we conducted in *C. elegans* are among the first efforts to systematically characterize genes on HSA21. Around the time of our loss-of-function screen, a systematic overexpression screen was conducted using HSA21 cDNAs in zebrafish embryos to identify developmental mutants (Edie et al., 2018). The work from the *C. elegans* and zebrafish studies has provided insight into gene function and identified genes that can be prioritized in future DS research.

Materials and Methods

Some descriptions that are relevant to the loss-of-function screen have been adapted from Nordquist et al., 2018.

Worm maintenance

C. elegans nematodes were kept at room temperature (~20℃) on an NGM (nematode growth media) agar plate seeded with an *E. coli* OP50 bacterial lawn. The N2 Bristol strain was used as the reference wild-type strain. Plates were regularly propagated to prevent starvation.

Identification of Orthologs

There are around 200 protein-coding genes on HSA21, and 165 genes remain after removing those that encode keratin. The initial list of orthologs was created using OrthoList, which reported 85 putative orthologs. Candidate orthologs were analyzed using InParanoid algorhythms, which predicts the human-to-worm relationship. All putative orthologs were considered for investigation except those within many-to-many relationships (many potential human genes corresponding to many potential *C. elegans* genes).

Loss-of-Function Screen

Ultimately, we investigated 51 tightly-conserved *C. elegans* genes that correspond to 47 human genes. Loss-of-function mutant strains were obtained through the National BioResource

Project (NBRP) and Caenorhabditis Genetics Center (CGC). Mutants underwent the following assays (described below): long-term locomotion, short-term locomotion, pharyngeal pumping, aldicarb sensitivity, and levamisole sensitivity. All behavioral assays were conducted blind to the worms' genotype.

Overexpression Screen

Orthologs are cloned into plasmids and injected directly into eggs of adult worms. The plasmids are expressed as extrachromosomal arrays and contain an mCherry reporter as the co-injection marker. F1 worms are isolated and propagated for experimentation if they reliably express the red marker. To date, 30 overexpression strains are available, covering 9 of the HSA21 orthologs injected at low and intermediate concentration and 8 control strains that only contain the co-injection marker. Long-term locomotion and egg retention assays have been performed (described below). Verification of transcriptional overexpression is conducted using qRT-PCR.

Long-term Locomotion / Exploration Assay

A single L4-stage worm is placed on a 3.5-cm diameter plate seeded with a thin, uniform bacterial lawn. The worm freely roams the plate for a 16-hour period. The tracks left in the lawn are compared against a grid system to measure the number of squares entered by the worm. This assay provides a read-out not only of general motor ability but also sensorimotor integration.

Short-term Locomotion / Radial Dispersion Assay

Five to eight worms are placed in the center of a 10-cm diameter plate seeded with a uniform bacterial lawn and allowed to disperse for 10 minutes before displacement is measured. This assay evaluates general motor abilities.

Pharyngeal Pumping Assay

Pharyngeal pumping rate of day one-staged adult worms was observed by counting the number of pumps in a 30-second time period.

Aldicarb Sensitivity Assay

Aldicarb is an acetylcholinesterase inhibitor that prevents the breakdown of acetylcholine in the synapses. In the presence of aldicarb, acetylcholine is not degraded and is thus able to constitutively activate the cholinergic receptors in muscle, causing paralysis. Aldicarb sensitivity was measured as described (Mahoney et al., 2006). Worms were placed on a 1-mM aldicarb plate and scored for paralysis every half hour for 3 hours. Worms were considered paralyzed when they did not spontaneously move and did not respond to prodding.

Levamisole Sensitivity Assay

Levamisole is a cholinergic receptor agonist. It constitutively activates the cholinergic receptors in muscle, causing paralysis if the receptors are working properly. Levamisole sensitivity was measured as described (Lewis et al., 1980). Day-one adult worms were placed on 800 μM Levamisole plates and scored for paralysis every 10 minutes for 1 hour.

Egg Retention Assay

A single adult worm is placed in a 20% bleach solution. This concentration dissolves the adult but leaves intact the eggs, which are then counted. Proper egg-laying requires regulated coordination between neuronal and muscular elements near the vulva.

qRT-PCR

Transcriptional overexpression of orthologs is assessed using qRT-PCR. RNA is isolated from crowded plates using a freeze-thaw technique. RNA is extracted using the Invitrogen RNA extraction kit and protocol. Following purification, I conduct reverse transcriptase (RT)-PCR using the ProtoScript II First Strand cDNA Synthesis Kit standard protocol. I conduct qPCR using SYBR Green according to the two-step protocol by Applied Biosystems. I optimized primer concentrations at 3.0 μL of 5 μM for both primers. Relative fold gene expression is quantified using the $\Delta\Delta C_T$ method. The housekeeping gene *pmp-3* is used as the reference gene. Each gene expression quantification required at least 3 technical and 3 biological replicates per primer and background.

Results

Loss-of-function Screen

Figure 1. Results from the short-term locomotion assay in loss-of-function mutants.

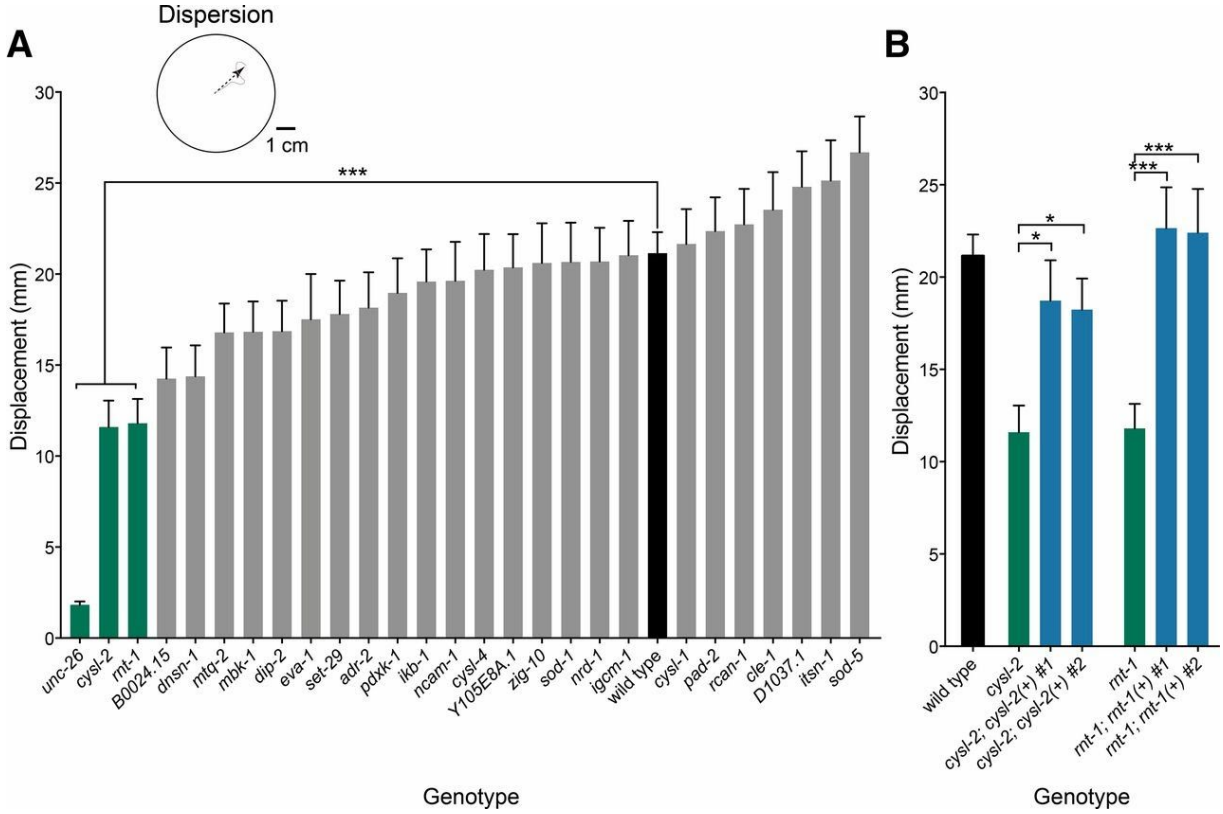

Figure 1. Results from the short-term locomotion assay in loss-of-function mutants. A. Displacement in mm by mutant genotype relative to wild-type N2. Three genotypes resulted in significantly reduced displacement: unc-26, cysl-2, mt-1. B. Displacement in mm by mutant genotype or restored genotype. The motor defect was rescued for both cysl-2 and mt-1 when given a functional copy of the ortholog. These data are modified from Figure 2 in Nordquist et al., 2018.

Figure 2. Results from the long-term locomotion assay in loss-of-function mutants.

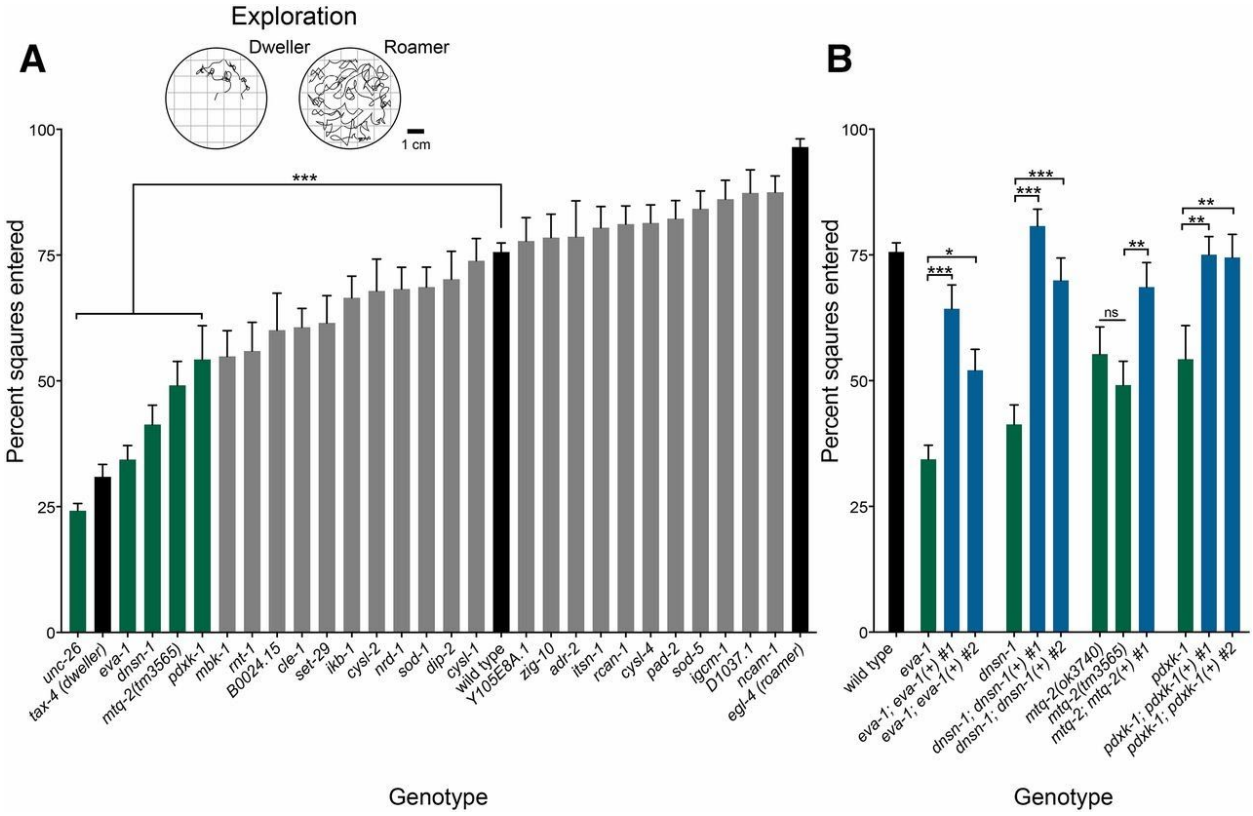

Figure 2. Results from the long-term locomotion assay in loss-of-function mutants. A. Percent squares entered by mutant genotype. Many genotypes were “dwellers” relative to wild-type, near the “dweller” control tax-4: unc-26, eva-1, dnsn-1, mtq-2, pdxk-1. B. Percent squares entered by mutant genotype or restored genotype. The motor defect was rescued in at least one restorative genotype for each mutant. These data are modified from Figure 3 in Nordquist et al., 2018.

Figure 3. Results from pharyngeal pumping assay for loss-of-function mutants.

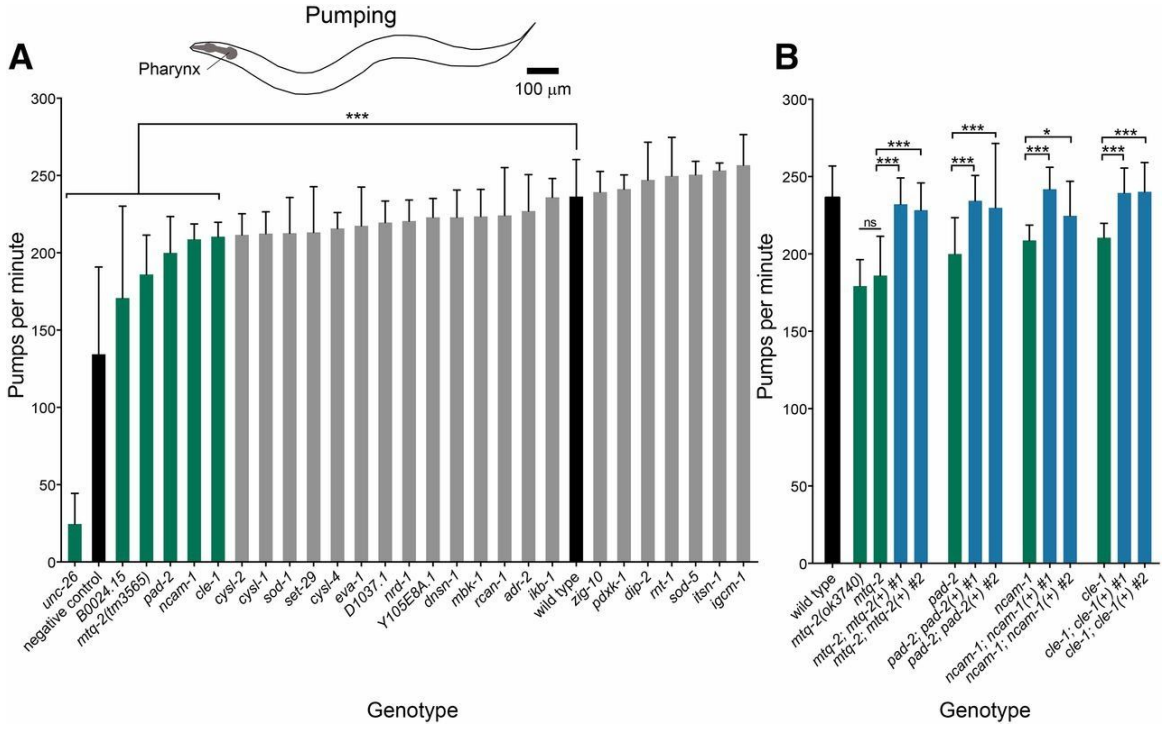

Figure 3. Results from pharyngeal pumping assay for loss-of-function mutants. A. Pumps per minute by mutant genotype. Several genotypes have a significantly reduced pumping rate: unc-26, B0024.15, mtq-2, pad-2, ncam-1, and cle-1. B. Pumps per minute by Genotypes mtq-2, pad-2, ncam-1, and cle-1. Each of the restorative genotypes resulted in rescuing the defective pumping. These data are modified from Figure 4 in Nordquist et al., 2018.

Figure 4. Results from the aldicarb assay for loss-of-function mutants.

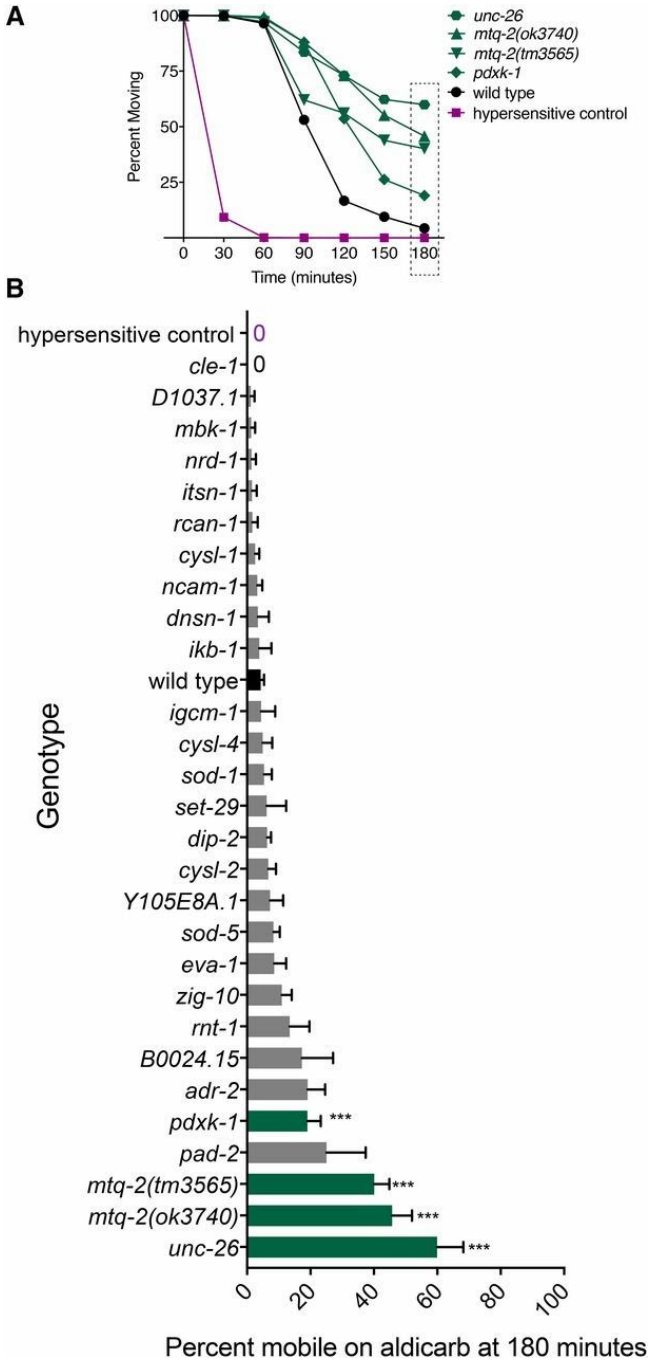

Figure 4. Results from the aldicarb assay for loss-of-function mutants. A. Percent moving on aldicarb plates over time. Genotypes unc-26, mtq-2 (ok3740 and tm3565 alleles), and pdxk-1 mutants showed significant resistance to paralysis relative to wild-type. B. Percent mobile on aldicarb at the 180-minute time point by genotype. There is significant resistance to paralysis in pdxk-1, mtq-w, and unc-26 genotypes. These data are modified from Figure 5 in Nordquist et al., 2018.

Figure 5. Results for aldicarb and levamisole assays for mtq-2, pdxk-1, and unc-26 mutant and restored genotypes.

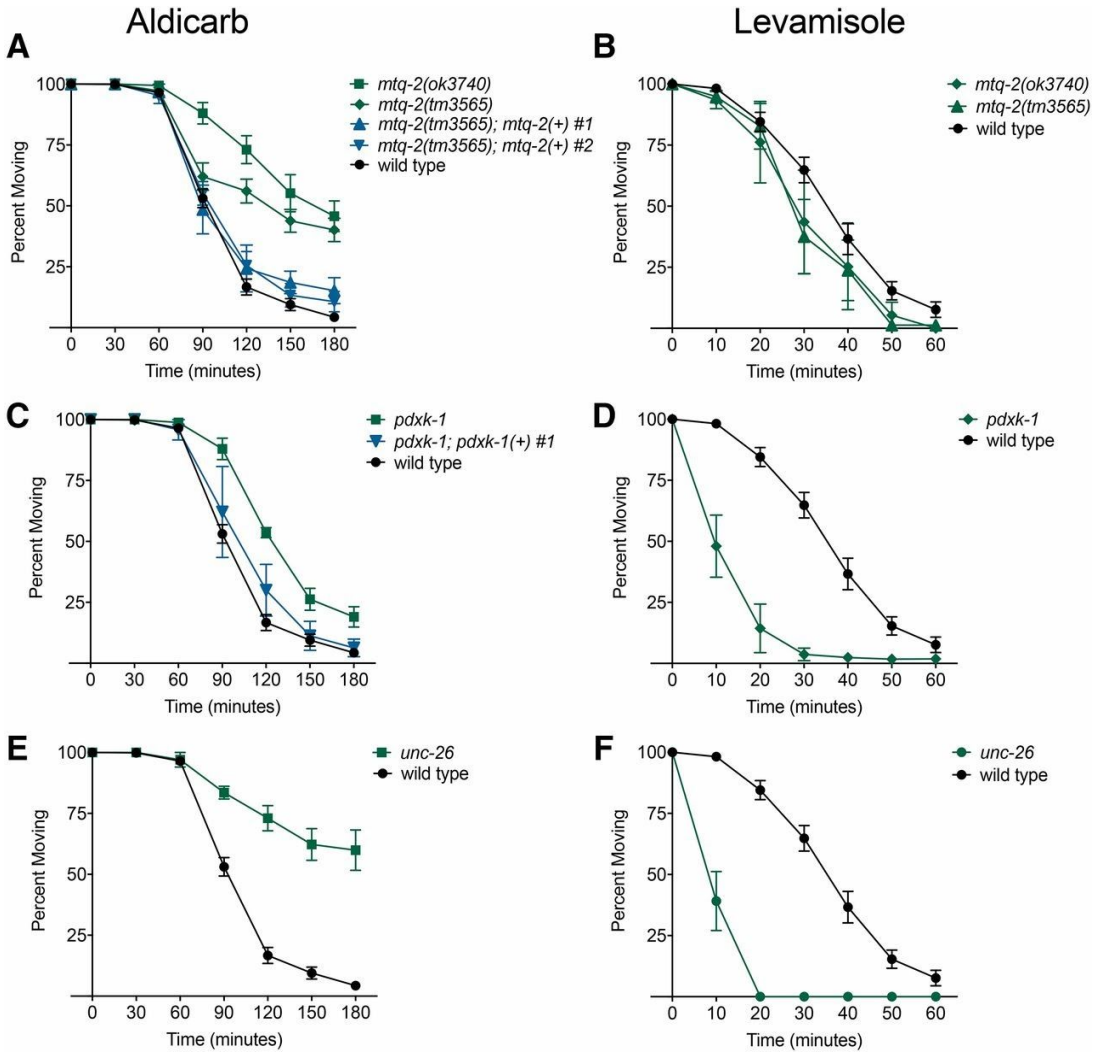

Figure 5. Results for aldicarb and levamisole assays for mtq-2, pdxk-1, and unc-26 mutant and restored genotypes. A. Aldicarb results for mtq-2 mutant. Green lines indicate mtq-2 mutants, which were resistant to paralysis by aldicarb. Blue indicates the restoration of the tm3565 allele, which rescued the sensitivity back to wild-type levels. B. Levamisole results for mtq-2. Mtq-2 mutants are sensitive to paralysis by levamisole. C. Aldicarb results for pdxk-1. Pdxk-1 mutants were resistant to paralysis by aldicarb, and this defect was rescued to wild-type levels. D. Pdxk-1 mutants are hyper-sensitive to paralysis by levamisole. E. Unc-26 mutants are resistant to paralysis by aldicarb. F. Unc-26 mutants are hyper-sensitive to paralysis by levamisole. These data are modified from Figure 6 in Nordquist et al., 2018.

Overexpression screen

Figures 6 and 7. Current results for behavioral assays for C. elegans overexpressing HSA21 orthologs.

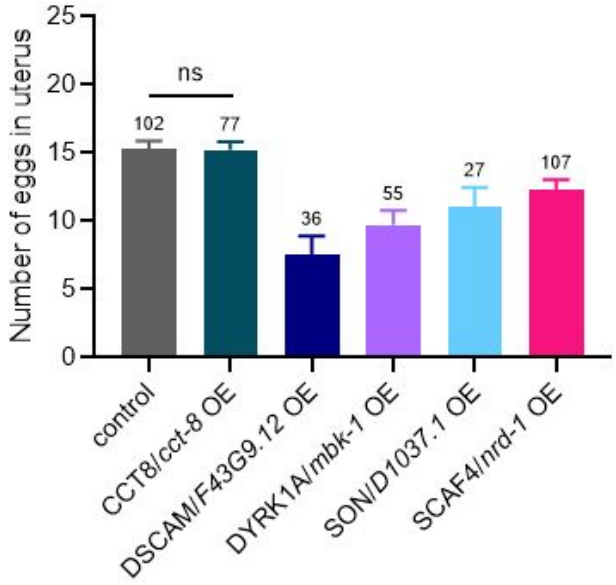

Figure 6. Day 1 Egg Retention

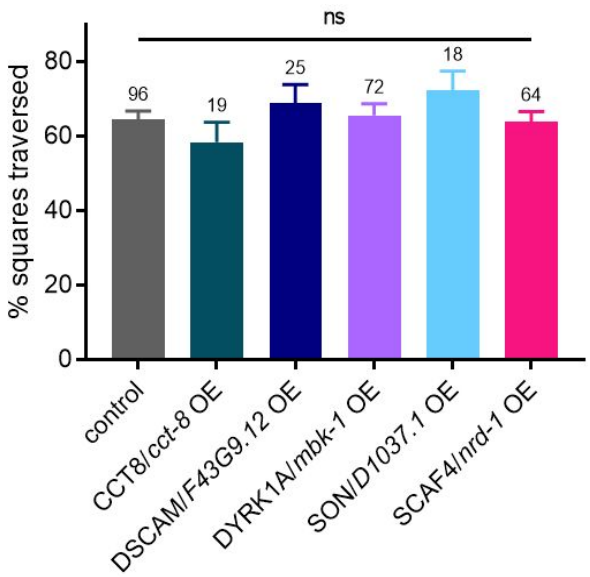

Figure 7. Long-term Locomotion

Figures 6 and 7. Current results for behavioral assays for *C. elegans* overexpressing HSA21 orthologs. 6. Egg retention in overexpression strains. Four out of five strains have a significantly low number of eggs retained relative to wild-type. 7. Long-term locomotion in overexpression strains. So far, none of the overexpression mutants vary significantly from wild-type in long-term locomotive ability.

Figure 8. Determining co-localization of MTQ-2 in synapses.

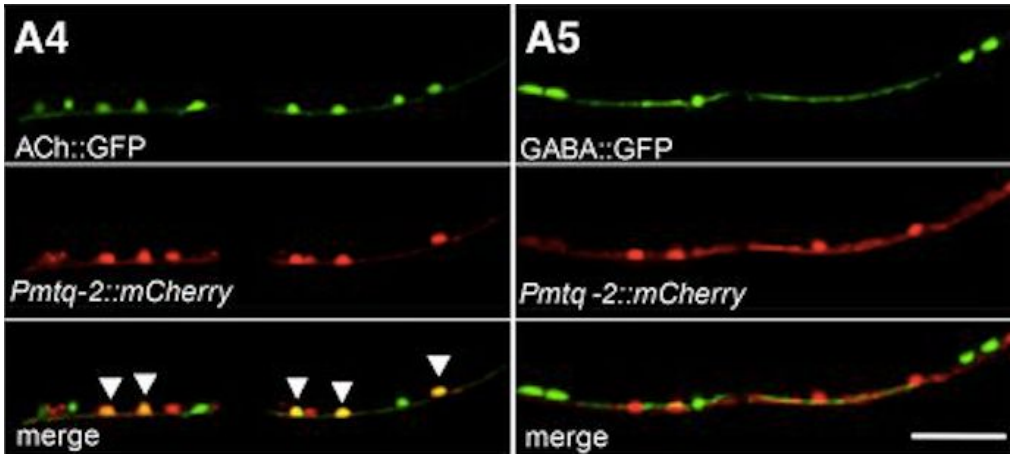

Figure 8. Determining co-localization of MTQ-2 at synapses. A. MTQ-2 colocalizes with acetylcholinergic synapses. Pmtq-2::mCherry colocalizes with ACh::GFP expression. B. MTQ-2 does not colocalize with GABAergic synapses. Pmtq-2::mCherry does not colocalize with GABA::GFP. These data are modified from Figure 7 in Nordquist et al., 2018.

Figure 9. Preliminary results for transcriptional overexpression from qRT-PCR.

Ortholog overexpressed	Relative to N2 Background	Relative to C5 Background
Cct-8 - High concentration	415.4	36.7
Cct-8 - Medium concentration	75.4	6.7
Cct-8 - Low concentration	28.8	2.5
Mtq-2 - Low concentration	26.8	165.5
Mbk-1 - Medium concentration	409.2	

Figure 9. Preliminary results for transcriptional overexpression from qRT-PCR. These data are generated using the $\Delta\Delta C_T$ method for relative quantification. Transcriptional overexpression is confirmed in cct-8, mtq-2, and mbk-1 overexpression strains relative to both N2 and C5 control backgrounds using pmp-3 as the reference gene. Relative overexpression varies widely but indicates successful overexpression of the transgenes.

Discussion

Our research represents one of the earliest attempts to characterize individual genes on the human 21st chromosome. We have identified several candidate genes that could contribute to DS-associated phenotypes by examining the orthologs in *C. elegans.* Our loss-of-function screen identified 10 genes that are critical for neuron or muscle function: *ncam-1, eva-1, cle-1, unc-26, cysl-2*, *rnt-1, pad-2, dnsn-1, pdxk-1,* and *mtq-2.* Critical mutants were identified by defects in the behavioral assays (Fig. 1-3) and the pharmacological assays (Fig. 4-5), and our work was published in *G3: Genes, Genomes, Genetics* (Nordquist et al., 2018). The results of the screen confirmed deficits for well-studied genes but also revealed roles for previously unstudied genes in neuron or muscle function. Performing a broad, unbiased screen of all available HSA21 orthologs enabled us to implicate previously unstudied genes in contributing to DS-associated phenotypes.

Figures 4 and 5 show the results for the aldicarb and levamisole assays. Aldicarb, the acetylcholinesterase inhibitor, prevents degradation of acetylcholine in synapses and leads to paralysis by over-stimulation of muscle. Three of the mutant genotypes (*mtq-2, unc-26,* and *pdxk-1*) were resistant to paralysis by aldicarb, with worms maintaining mobility on aldicarb plates longer than wild-type. This either indicates that mutants have less acetylcholine released into synapses or have an issue with receptor binding to acetylcholine. The levamisole assay was conducted to clarify whether the mutants had pre- or post-synaptic defects. Levamisole is a cholinergic receptor agonist, so it triggers paralysis by over-stimulating muscle when the postsynaptic receptors work properly. The three mutants were all found to be sensitive or hypersensitive to paralysis by Levamisole,

indicating that resistance to aldicarb is caused by defects in presynaptic neurotransmitter release. Further investigation of *mtq-2* verified that it is expressed throughout the worm's nervous system and colocalized to cholinergic synapses (Fig. 8), which suggests its promise as a potential contributor to neuromuscular symptoms experienced in DS.

Preliminary results from the overexpression screen are shown in Figures 6 and 7. Of the 5 strains examined thus far, 4 display significant reduction in egg retention and none display defects in long-term locomotion. Fewer retained eggs could indicate hyperactive neurotransmission within the egg-laying pathway, or could be due to developmental defects that lead to reduced or delayed egg production. We plan to assess both egg-laying behavior and gross developmental defects moving forward. To verify the overexpression of the transgenic arrays, I began conducting qRT-PCR on overexpression strains relative to N2 wild-type background and the C5 control background that expresses the co-injection marker mCherry. Relative overexpression of each target gene varies significantly, but we satisfied our goal of confirming the overexpression of the desired genes.

We have identified genes that are critical for neuromuscular function through the loss-of-function screen, and now we aim to identify the orthologs that are problematic when overexpressed. These genes are promising candidates for driving DS-associated phenotypes like muscle weakness and intellectual disability. Our work in *C. elegans* examines the genetic basis for DS-associated phenotypes at the level of individual genes, which provides a necessary complement to the foundational studies in other model systems. By characterizing the functions of candidate genes, we hope to identify potential therapeutic targets for alleviating DS-associated symptoms.

Acknowledgements

From Nordquist et al., 2018: "We thank the *Caenorhabditis* Genetic Center [funded by the National Institutes of Health (NIH)] and National Bioresource Project, Yishi Jin, and Harald Hutter for strains. Susan Rozmiarek provided expert assistance, Allison Griffith provided preliminary data, Luisa Scott provided constructive input, and Sophie Sanchez provided editing. Research was inspired by Ocean Pierce-Shimomura. Funds were provided by the Alzheimer's Association, Down Syndrome Research and Treatment Foundation/LuMind, Research Down Syndrome, Jerry and Judy Horton with the Point Rider Foundation, and NIH T-R01 awards 1R01AG041135 and 1RF1AG057355. The authors declare no conflict of interest."

Many of the assays conducted in the loss-of-function screen were conducted by Dr. Sarah Nordquist. I performed all long-term locomotion assays and assisted Sarah with tasks like worm maintenance, performing PCRs, optimizing primers, etc. so that assays could be conducted more efficiently. I also performed many behavioral assays (long-term locomotion, short-term locomotion, and egg retention) in the pilot stage of the overexpression screen. I conducted all qPCR experiments including all cDNA processing steps.

I want to thank Dr. Jon Pierce for his encouragement and guidance throughout my time in the lab. He was always enthusiastic and willing to help when I needed to consult him for advice. He and lab manager Susan Rozmiarek helped me secure funding sources for summer research fellowships. Susan also helped me keep my projects organized, and was a friend and advocate. My graduate student mentor Dr. Sarah Nordquist trained me in essential techniques and entrusted me with my first independent part of the project. My other graduate mentor Sophie Sanchez talked with me regularly about the project, and I really enjoyed exchanging ideas with her. Sophie provided valuable help as I began piloting qPCR. Both of my graduate student mentors were encouraging, kind, and approachable for any questions I had. The Pierce lab has been a wonderfully supportive environment for me and sets a solid foundation for my future academic work.

References

1. Antonarakis, Stylianos E., et al. "Down syndrome." *Nature Reviews Disease Primers* 6.1 (2020): 1-20.
2. Belichenko, N. P., Belichenko, P. V., Kleschevnikov, A. M., Salehi, A., Reeves, R. H., & Mobley, W. C. (2009). The "Down syndrome critical region" is sufficient in the mouse model to confer behavioral, neurophysiological, and synaptic phenotypes characteristic of Down syndrome. *Journal of Neuroscience*, *29*(18), 5938-5948.
3. Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics*, *77*(1), 71-94.
4. Edie, Sarah, et al. "Survey of human chromosome 21 gene expression effects on early development in Danio rerio." *G3: Genes, Genomes, Genetics* 8.7 (2018): 2215-2223.

5. Kazuki, Yasuhiro, et al. "A non-mosaic humanized mouse model of Down syndrome, trisomy of a nearly complete long arm of human chromosome 21 in mouse chromosome background." *bioRxiv* (2019): 862433.
6. Lewis, J. A., Wu, C. H., Berg, H., & Levine, J. H. (1980). The genetics of levamisole resistance in the nematode Caenorhabditis elegans. Genetics, 95(4), 905-928.
7. Lyle, Robert, et al. "Genotype–phenotype correlations in Down syndrome identified by array CGH in 30 cases of partial trisomy and partial monosomy chromosome 21." *European Journal of Human Genetics* 17.4 (2009): 454-466.
8. Mahoney, T. R., Luo, S., & Nonet, M. L. (2006). Analysis of synaptic transmission in Caenorhabditis elegans using an aldicarb-sensitivity assay. *Nature protocols*, *1*(4), 1772-1777.
9. Nordquist, S. K., **Smith, S. R**., & Pierce, J. T. (2018). Systematic functional characterization of human 21st chromosome orthologs in Caenorhabditis elegans. *G3: Genes, Genomes, Genetics*, *8*(3), 967-979.
10. Olson, L. E., Roper, R. J., Sengstaken, C. L., Peterson, E. A., Aquino, V., Galdzicki, Z., ... & Reeves, R. H. (2007). Trisomy for the Down syndrome 'critical region' is necessary but not sufficient for brain phenotypes of trisomic mice. *Human molecular genetics*, *16*(7), 774-782.
11. Patterson, D. (2009). Molecular genetic analysis of Down syndrome. *Human genetics*, *126*(1), 195-214.
12. Pelleri, M. C., Cicchini, E., Petersen, M. B., Tranebjærg, L., Mattina, T., Magini, P., ... & Seri, M. (2019). Partial trisomy 21 map: Ten cases further supporting the highly restricted Down syndrome critical region (HR-DSCR) on human chromosome 21. Molecular genetics & genomic medicine, 7(8), e797.