Genetic regulators of stress-induced RNA mis-splicing in *Caenorhabditis elegans*

A Thesis Submitted to the
College of Graduate and Postdoctoral Studies
in Partial Fulfillment of the Requirements
for a Master of Science Degree
in the Department of Veterinary Biomedical Sciences
University of Saskatchewan
Saskatoon

By

SAMANTHA CORINNE CHOMYSHEN

Permission to use

In presenting this thesis/dissertation in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis/dissertation in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis/dissertation work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis/dissertation or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis/dissertation.

Requests for permission to copy or to make other uses of materials in this thesis/dissertation in whole or part should be addressed to:

Head of the Department of Veterinary Biomedical Sciences

52 Campus Drive

University of Saskatchewan

Saskatoon, Saskatchewan S7N 5B4 Canada

OR

Dean

College of Graduate and Postdoctoral Studies

University of Saskatchewan

116 Thorvaldson Building, 110 Science Place

Saskatoon, Saskatchewan S7N 5C9 Canada

Abstract

Splicing of pre-mRNA is an essential process for all eukaryotic dividing cells. Pre-mRNA splicing defects are implicated in numerous human diseases, including Alzheimer's disease and cancer, however its cause is poorly understood. Using the nematode *Caenorhabditis elegans* as a model, the Wu lab has recently shown that exposure to the environmental heavy metal cadmium can cause RNA splicing disruption, implicating loss of RNA metabolism regulation as a potential mechanism of cadmium toxicity. To understand the genetic mechanism of RNA splicing regulation under environmental stress, I sought to identify and characterize genes that, when knocked down, can protect against RNA splicing errors. Using a *C. elegans in vivo* splicing reporter, I found that an overwhelming majority of the gene knock-downs that improved RNA splicing under stress encode various components of the translation machinery, including *ifg-1*, which encodes the human eIF4G gene previously shown to regulate aging. Knockdown of various protein translation related genes has been shown to not only increase *C. elegans* lifespan but now also to enhance resistance to cadmium survival. Using RNA-sequencing, I found that *ifg-1* partial loss of function mutants show increases in expression of >80 genes that regulate RNA splicing; importantly, *ifg-1* mutants exposed to cadmium show a 50% decrease in cadmium-induced alternative splicing events observed in wild-type worms. Downstream of *ifg-1*, I have identified the SMA family of transcription factors as key regulators that are required for RNA splicing protection under stress in the *ifg-1* mutants. Suppression of translation has previously been shown to be beneficial in promoting longevity and stress resistance in various organisms including *C. elegans*, and my study may have implicated a novel mechanism through which these physiological benefits are achieved in part by improvements to RNA splicing fidelity.

Acknowledgements

I would like to take this opportunity to first and foremost thank my supervisor Dr. Wu for not only giving me the incredible opportunity to study this field of research, but also for the vast amount of support throughout the past two years. As one of his first students, he presented me with the opportunity to perform research in an ever-growing field, as well as the opportunity to watch his lab grow from three people with two microscopes to the small family we are now. I also am in debt to the enormous amounts of help given to me by all members of the Wu lab, past and present; Sydney Murray for her guidance at the beginning of my research, Brandon Waddell for his lessons and help throughout the entire time, and each and every other lab member that has come through the Wu lab since the beginning. Without your help, my research would not be where it is today. Beyond the lab, I would like to thank my group of friends for the weekly entertainment through a pandemic Master's program, especially to my best friend Robyn (b) for always being there for me. I would also like to thank my sources of funding for making this research possible; my work was funded by an NSERC Discovery grant, a SHRF Establishment Grant, a CFI Infrastructure grant, and a Graduate Teaching Fellowship provided to me by the Department of Veterinary Biomedical Sciences.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

A	Adenine
ALS	Amyotrophic lateral sclerosis
AS	Alternative splicing
ATP	Adenine triphosphate
ATP6AP2	ATPase H+ Transporting Accessory Protein 2
BMS1	Ribosome biogenesis factor 1
bp	Base pair
C. elegans	*Caenorhabditis elegans*
cDNA	Complimentary DNA
cdr-1	Cadmium responsive 1
ced-1	Cell death abnormality 1
circRNA	Circular RNA
CRISPR	Clustered regularly interspaced short palindromic repeats
csp-2	Caspase 2
CT	Cycle threshold
daf-16	Dauer formation 16
DAVID	Database for Annotation Visualization and Integrated Discovery
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
dsRNA	Double stranded RNA
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid

egl-45	Egg laying defective 45
eif	Eukaryotic initiation factor
eIF4G3	Eukaryotic Translation Initiation Factor 4 Gamma 3
EV	Empty vector
FDR	False Discovery Rate
FOXO	Forkhead family of transcription factors
FPKM	Fragments Per Kilobase of transcript per Million mapped reads
G	Guanine
GC Content	Guanine-Cytosine content
GFP	Green fluorescent protein
hrp-2	Heterogeneous nuclear ribonucleoprotein R homolog
hrpf-1	HnRNP F homolog
hsp-16.49	Heat shock protein 16.49
ifg-1	Initiation Factor 4G family 1
IncLevel	Inclusion level
inf-1	Initiation factor 1
IPTG	Isopropylthio-β-galactoside
LB	Lysogeny broth
MAD	Mothers against decapentaplegic
mRNA	Messenger RNA
NER	Nucleotide excision repair
NGM	Nematode growth media
nhr-61	Nuclear hormone receptor 1

numr-1	Nuclear localized metal responsive 1
p value	Pearson correlation coefficient
PCR	Polymerase chain reaction
qPCR	Quantitative polymerase chain reaction
ret-1	Reticulon protein 1
RFP	Red fluorescent protein
RNA	Ribonucleic acid
RNAi	RNA interference
ROX	Carboxyrhodamine
rpl	Ribosomal protein, large subunit
rps	Ribosomal protein, small subunit
rsp-2	SR Protein 2
sfa-1	Splicing factor 1
siRNA	Small interfering RNA
SMA	Small worm phenotype
SMAD	Small worm phenotype + mothers against decapentaplegic
snr/ snRNP	Small nuclear ribonucleoprotein
SRRM2	Serine/Arginine Repetitive Matrix 2
TAE	Tris base, acetic acid and EDTA
TGF-β	Transforming growth factor-β
TOR	Target of rapamycin
tRNA	Transfer ribonucleic acid
U	Uracil

uaf-2	U2AF splicing factor

1. Introduction

1.1: RNA splicing

From DNA to RNA to protein – the central dogma of biology defines the molecular pathway in which a cell produces a protein from a strand of DNA. This basic process is evolutionarily conserved across the taxonomic domain of Eukarya, with minor kingdom-specific variations. All eukaryotes, from the smallest bacteria to the largest plant, undergo this highly regulated three-step process in order to sustain life. A signal is first sent to the nucleus of a cell to separate a double strand of DNA into single strands, which are used as template code for producing messenger RNA (mRNA); this is known as DNA transcription. As the product is an exact copy of the DNA strand, this results in replication of a DNA strand in the form of single stranded RNA desired by the cell in its current state. After transcription is complete, post-transcriptional modifications at the 5' or 3' end must take place to transform the immature precursor messenger RNA (pre-mRNA) into a proper code for the desired protein. In 5' processing, the strand of pre-mRNA is capped by a 7-methylguanosine to protect from ribonuclease degradation. In 3' modification, the pre-mRNA can be cleaved and polyadenylated to protect against ribonuclease degradation, processed by histones to protect against damage, or spliced via RNA splicing to create a final mRNA for protein translation. Once modified, the final step takes place – protein translation. The mature mRNA is used as a code for transfer RNA (tRNA) to be bound together to form a protein. This entire process is tightly regulated – even a single missing base pair will cause the strand of mRNA to produce a completely nonsensical protein to be transcribed. However, even under the tightest regulation, errors can occur; especially in post-transcriptional modification.

The pre-mRNA splicing process is highly prevalent in eukaryotes and less common amongst prokaryotes. In general, a strand of pre-mRNA is composed of two different types of sections of nucleotides: exons and introns. In RNA splicing, pre-mRNA is excised and joined together in order to remove the long non-coding introns from the strand and join together the coding exons. To do this, a series of up to 300 small nuclear riboproteins (snRNPs) form a complex known as the spliceosome [1]. While each of these proteins has a distinct role in RNA splicing, there are five major proteins regarded as the main splicing proteins: the U1, U2, U4, U5, and U6 snRNPs. To begin, the mRNA is cleaved by the U1 snRNP at a pre-determined

splice site at the 3’ end of an intron; while minor variations can occur, it is highly conserved that the splice sites contain a GU at the 5’ end and an AG at the 3’ end, both of which regions are flanked by long strands of nucleotides with low sequence homology to the splice site. The spliced end is then brought towards a branch point region typically 18 to 40 nucleotides downstream where it is attached in a lariat configuration to prevent re-attachment to the splice site as well as RNA degradation. After this, the lariat and the cut end of the pre-mRNA are guided to the 5’ end by the U2 snRNP and the snRNP complex U4/U6; U5 then assists in excising the splice site at the 5’ end of the intron, as well as binding the two ends of the pre-mRNA strand together. Once the pre-mRNA is covalently bound, the removed intron lariat is released with U2, U5, and U6 still bound, which will be degraded and the snRNPs reused for further RNA splicing. This process is repeated multiple times along the strand of pre-mRNA at sites directed by the cell to create the desired mRNA molecule that will be translated into a final protein – all while transcription, other post-translational modifications, and even nuclear export are occurring.

The existence of RNA splicing alone makes little sense – why do long sections of non-coding RNA exist if their only purpose is to be spliced out? Besides acting as an intermediate product for other post-translational modifications to take place, what purpose does pre-mRNA have? These questions were brought up alongside the identification of exons and introns [2], and their answer first theorized one year later – different sections in a pre-mRNA strand may be spliced out to create different final proteins [3]. Today, this process is known as alternative splicing (AS) – the ability to create different mRNA molecules from a single strand of pre-mRNA. The amount of pre-mRNA needed in a cell can be vastly reduced by AS; while there may only be ~20,000 protein-coding genes in the human genome, it is estimated that the human proteome contains 620,000 to 6.13 million protein species, a diversity that is created by AS [4], [5]. Through bioinformatics analysis of human mRNA-seq data, Pan et al. estimates that all multi-exon genes in the human genome undergo at least one AS event, with up to 100,000 high-abundance events [6].

To understand the needs of a cell, the nucleus utilizes a series of transcription factors that retrieve information from the interior and exterior environment of the cell. These signals convey information about the general state around the cell, as well as any changes within the cell itself,

into the nucleus. Transcription factors are generally dormant, only activated once external stimuli arrives. Once stimulated, they can either signal through the nuclear wall or translocate into the nucleus to convey their signal [7]. While there are a large variety of transcription factors in humans, one of note is the SMAD family of transcription factors, whose acronym is derived from the *Caenorhabditis elegans* "small worm phenotype" (SMA) family of proteins and the *Drosophila* "mothers against decapentaplegic" (MAD) family of proteins. The SMAD proteins respond to signals produced from the cell wall dwelling transforming growth factor-β (TGF-β) cytokines, conveying the information produced by these growth factors to the nucleus of the cell. The SMAD proteins specifically transport signals to the nucleus that relate to the transcription of genes involved in growth and development, making them critical for regulating homeostasis of the organism [8]. Once activated, transcription factors act in the nucleus by binding to specific motif elements in DNA to initiate the transcription of genes to exert downstream effects.

Functionally coupled to transcription, the spliceosome acts on the pre-mRNA to create different final products. The final protein produced can be influenced by alternative splicing, which is determined by the current state of the cell – environmental conditions, developmental stages, even depending on the type of cell. A series of complex proteins signal to the nucleus to begin the process of creating the desired protein, at which point transcription and post-transcriptional modifications like AS can occur. There are five major types of AS characterized to date: exon skipping, alternative 3' or 5' splice sites, mutually exclusive exons, or intron retention (Figure 1.1). In exon skipping, the 5' splice site at the end of the intron being spliced is not used; instead, a 5' splice site further downstream is selected for resulting in the skipping of an exon. The intron lariat is carried downstream to the desired splice site, then the lariat and all exons and introns included in it are released, achieving in skipping over an exon all together. In alternative 3' or 5' splice sites, a different splice site within an exon is used at the beginning or the end of an mRNA splicing event, allowing for a different start or end of that exon. Mutually exclusive exons sees the inclusion of one exon for one protein and a second exon for a second protein, but the two exons are never present in both proteins. And finally, as the name suggests, intron retention results in an intron that does not get spliced out and is instead retained in the mature mRNA strand. Each of these different events allows for a high degree of protein diversity to be produced from a cell without the inclusion of a different gene for each final protein.

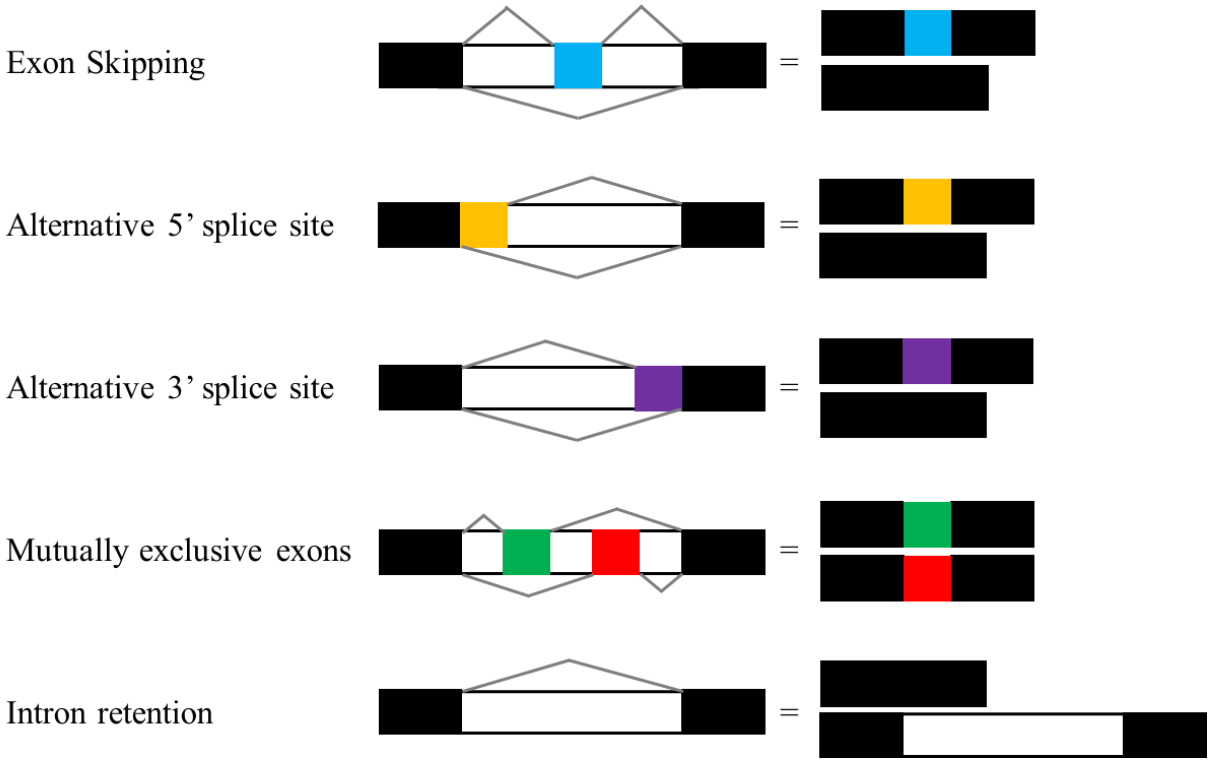


Figure 1.1. Schematic of the five types of AS. From top to bottom, they are: exon skipping, alternative 5’ splice site, alternative 3’ splice site, mutually exclusive exons, and intron retention. White boxes represent introns, all other boxes represent exons, and diagonal grey lines represent the path a transcript may take. The two possible transcripts produced are displayed to the right of each event.

1.2: RNA mis-splicing

As discussed in the previous section, AS is an essential function of the cell necessary to produce mRNA that will be translated into distinct proteins while reducing the amount of different genes a cell has to encode. However, this process can be disrupted in cells during times of stress; this is known as RNA mis-splicing. By inducing AS without proper signalling, mRNA can be produced that encodes either the wrong protein or a non-functional protein aggregate. As well, RNA mis-splicing can occur via defects or mutations in the spliceosome complex itself [9]. Discussed as early as 1993, RNA mis-splicing is known to lead to dangerous protein aggregates of stable circular RNA (circRNA) products with no apparent function [10]. The cause of RNA mis-splicing can be as simple as a small mistake made by the spliceosome, or as complex as environmental shifts and genetic predisposition. On the molecular level, RNA mis-splicing has four theorized reasons for why it takes place, as described by Scotti and Swanson. First, this machinery relies on searching a strand of pre-mRNA and identifying the correct 5' and 3' splice sites for each RNA splicing event – a difficult task made harder by the long non-coding introns and made impossible by single base pair mutations in transcription. Second, knowing which protein the complex is splicing for is a poorly signalled task that is made harder by the preference to splice the alternative protein during embryonic and fetal development. Third, as mentioned in the first theory, introns are long coding sequences; when partnered with exon sequences that are short in comparison, it creates an easy opportunity for the spliceosome complex to miss the 5' splice all together and continue on to the 5' splice site of the next exon. And finally, AS is coupled to translation itself and therefore regulated by the RNA polymerase that controls transcription; if the desired 5' splice site is not yet transcribed, the spliceosome may pick a weaker, incorrect 5' splice site in a "first come first serve" model [11] of site recognition [9].

As well as the cellular causes of RNA mis-splicing discussed above, there are environmental changes that can induce the deleterious process. Under normal environmental conditions RNA splicing and AS thrive, however environmental changes can result in RNA mis-splicing. Disruption of RNA splicing has been observed under environmental stress conditions such as heat shock [12], ethanol exposure [13], and most recently induced by the heavy metal cadmium [14]. While other environmental conditions that disrupt splicing exist beyond these three, there is a gap in literature on the other environmental stressors can affect RNA splicing,

especially compared to the wide variety of studied environmental conditions that disrupt translation [15]. A recent study has shown that cadmium can directly disrupt RNA splicing in a dose dependent manner [14]; as such, cadmium presents an interesting chemical tool for studying mechanisms of environment induced RNA mis-splicing. Cadmium exists in the environment, disrupted by human mining activity for battery use. As well, cadmium is present in cigarette and cigarette smoke. As such, cadmium presents most toxicity in industrial workers and those who smoke. Cadmium enacts its toxicity through mimicking calcium in the human body, a disguise easily achieved due to the +2 charge on both ions. This results in a series of toxic effects on the human body, including neurotoxicity and cancerous effects [14].

1.2.1: RNA splicing in disease

When RNA mis-splicing occurs, the product is marked for degradation via a cellular surveillance mechanism called nonsense mediated decay, and removed by cells with the only consequence being a minor loss of energy in forms of ATP. However, when RNA mis-splicing becomes systemic via environmental stress or internal dysregulation (i.e. mutated spliceosome), the cell cannot adequately remove the incorrect product at a fast enough rate to ensure the health of the cell. When this occurs, the cell is overwhelmed by RNA splicing errors, forming large protein aggregates in the cytoplasm and preventing the desired protein from being produced. As such, the proteins necessary to reduce RNA mis-splicing may not get translated, and the cell enters a continuous feedback loop of RNA mis-splicing and lack of repair. The destruction caused by systemic RNA mis-splicing is consequential and implicated in many well-known diseases in humans, namely Alzheimer's disease, Parkinson's disease, certain forms of cancer [9]. The dysregulation caused by errors in RNA splicing can lead to many different states of disrepair, depending on the cell type or the gene that is mis-spliced. As such, errors in RNA mis-splicing is hard to target therapeutically. Until recently, the role of environmental influences on the development of neurodegenerative diseases have not being well established. However, a movement gaining increasing traction in literature hypothesizes that exposure to environmental stressors that upset RNA splicing, such as the aforementioned heat shock or cadmium, may act as an external inducer on these disease states. As such, it is of interest to understand the poorly

researched mechanism of RNA mis-splicing in an effort to one day develop potential therapeutic strategies.

Protein aggregates are a hallmark of every neurodegenerative disease – accumulation of insoluble non-degradable proteins gathering in the cytoplasm preventing the cell from performing normal functions. In the case of Alzheimer's disease, the two main protein aggregates are β-amyloid and Tau. What is unknown, however, is how these protein masses form or how to treat them. Bai et al. took a large step in uncovering these unknowns, with an analysis of the 4,200+ proteins in cultured Alzheimer's disease afflicted brain cells. Surprisingly, they identified a large amount of U1 snRNP constituents aggregated in the cytoplasm of neuronal cell bodies [16]. This spliceosome component and key splicing gene usually resides in the nucleus, with no reason to migrate to the cytoplasm. However, Bai et al. noticed the large aggregate and hypothesized that this loss of U1 snRNP function is a key factor in causing Alzheimer's disease and its symptoms; losing the function of U1 snRNP in the nucleus results in the loss of proper RNA processing before translation can take place. This hypothesis has since been further confirmed by follow up studies [17], [18], indicating that RNA mis-splicing does indeed play a major part in genetic Alzheimer's disease.

Just as in Alzheimer's disease, Parkinson's disease has its own indicative protein aggregate - α-synuclein. As early as 1991, RNA mis-splicing was implicated in the cause of different forms of Parkinson's disease. Maroteaux and Scheller identified three differently spliced isoforms of rat α-synuclein in Lewy body protein aggregates, determining the diversity is due to AS [19]. Eight years later, D'Souza et al. identify three different genes that are alternatively spliced in Parkinson's disease, all of which are regulatory elements for RNA splicing [20]. Each gene has its alternate form upregulated in Parkinson's disease, causing failure of RNA splicing. In 2010, Shehadeh et al. identified an upregulation of the short isoform of serine/arginine repetitive matrix 2 (SRRM2) by 1.7 fold, with the long, normal isoform downregulated 0.4 fold [21]. While SRRM2 may not be one of the core splicing genes, it does influence the expression of protein isoforms in a cell. In 2013, Korvatska et al. describe RNA mis-splicing as the underlying cause of X-linked parkinsonism with spasticity. They found an increase in a single exon skipped isoform of ATP6AP2, which decreased the relative abundance of the desired isoform and produced an excessive amount of the insoluble isoform [22].

ATP6AP2 is typically an ATPase used in degradation and autophagy, however the exon skipped isoform cannot perform the task required. Unsurprisingly, the cellular pathway that requires ATP6AP2 is one that is well known to be disrupted in patients with Parkinson's disease. There are many different forms of Parkinson's disease each with their own cause, however with more and more research performed it is becoming increasingly clear that RNA mis-splicing plays a central part in many of them, if not all of these disease states.

The ever-growing field of cancer research has only known about the implications of RNA mis-splicing in causing cancer for the last decade. Yoshida et al. were the first to find that myelodysplastic cells predisposed to leukemia expressed alternatively spliced isoforms of at least five major splicing genes, causing a large disruption to the entire RNA splicing pathway [23]. Since then, research has only grown on the topic. To assist in understanding the massive scale of RNA mis-splicing in cancer, Jung et al. performed whole-genome analysis of 1,134 pan-cancer genomes, identifying 678 somatic intronic mutations affecting RNA splicing [24]. RNA mis-splicing in cancer can be as simple as overexpression of an anti-apoptotic isoform of a protein to the deregulation of the RNA splicing pathway entirely via upstream changes to the mechanism [25]. Perhaps most obnoxiously, cancer cells have been observed to "hijack" RNA splicing for its own advantages, including undergoing isoform switching to either evade cancer treatments [26] or to promote cellular proliferation. The gene PRPF6, a member of the normally tightly regulated U4/U5/U6 snRNP complex that performs the 5' splice and rejoin of the two excised ends, can assist in colorectal cancer proliferation by splicing genes associated with cancer growth [27]. Despite all these consequences, there is ongoing research that targets the spliceosome in a manner that assists in "fixing" RNA splicing in an effort to treat the afflicting cancer [28], [29].

There are many more diseases in humans and other mammals that are either caused by or made worse by RNA mis-splicing, such as Hutchinson–Gilford progeria syndrome, many forms of muscular dystrophy and atrophy, cardiovascular issues, amyotrophic lateral sclerosis (ALS), and dilated cardiomyopathy [9]. All of these disease states are influenced either by differential isoform expression caused by altered RNA splicing or changes to the RNA splicing pathway all together, reducing its fidelity overall. As such, RNA splicing reveals itself as a unique target; if one is able to repair the altered RNA splicing pathway, one could then treat or cure the disease state at hand. However, methods to treat RNA splicing errors are currently limited to

experimental tests in certain types of cancers due to the large field of unknowns within this research.

1.2.2: RNA splicing in aging

Many theories for how aging occurs in an organism and why it exists in the first place have come to light in the past century and a half. The first commonly accepted modern theory of aging was the “wear and tear” theory of aging, presented in 1882 by Dr. August Wiesmann. This theory states that the damage done to an organism over time accumulates to a point of no return, at which death takes place. Since then, other theories of aging have been discussed either as replacements for the previously researched theory or ones that work in tandem with it. Currently, there are five major theories of aging in modern science: the wear and tear theory, the rate of living theory [30], the cross-linking theory [31], the free-radicals theory [32], [33], and the somatic DNA damage theory [34]. While each theory has plenty of supporting research, currently the somatic DNA damage theory is most commonly accepted. In this theory, DNA constantly accumulates damage that is identified and repaired by a cell. However, as one ages, the amount of DNA damage occurring increases and the mechanisms of DNA damage repair begin to fail, accumulating a high quantity of damage that a causes irreversible damage and cell death.

Over the past decade, research on the effect of altered protein synthesis on the aging of an organism has started to arise [35]. As one of the basic functions of a cell, any changes to protein synthesis will affect the state of the cell overall. However, post-transcriptional modifications must come before translation can occur, meaning that changes in RNA splicing are just as dangerous, if not more dangerous, than errors in translation. The connection between RNA mis-splicing and aging has been known for since 1977, when Yannarell et al. saw a three fold decrease in mRNA released into the cytoplasm in the livers of older rats, theorizing that aging first begins with faults in the nucleus [36]. However only recently has RNA mis-splicing been implied in early aging [37]. Gruner et al. first determined that there is an accumulation of circRNA in aging mice, hypothesizing that the circRNA play a role in the nervous system [38]. As we now know, circRNA encompass a large part of the human genome with a variety of functions; however, they can also be the product of RNA mis-splicing, and can form insoluble

protein aggregates due to their stable end-free design. In the same year, Rodríguez et al. performed analysis on five different types of mouse tissue, finding an accumulation of alternatively spliced products in all five samples [39]. Not only this, they found that 158 of the differently spliced genes were involved in RNA processing, leading to their emphasis on the importance of an unknown AS mechanism in aging. There is an ample amount of research supporting RNA mis-splicing in aging-related diseases such as Alzheimer's disease and Parkinson's disease as mentioned earlier. However, amidst all the research available, there is little to no consensus on the mechanism of how RNA mis-splicing and aging are linked. A prominent question in this field of literature is a "chicken or the egg" type question; does abnormal RNA splicing lead to the onset of aging, or is RNA mis-splicing a product of age-related decline? If the former is correct, then RNA mis-splicing can be used as a target to delay the onset of aging or prevent aging-related diseases. And if the latter is correct, then reducing the accumulation of aging-related defects would allow for protection from RNA mis-splicing.

1.3: *Caenorhabditis elegans* as a model organism

While a single one millimeter long transparent nematode might not seem very special, *Caenorhabditis elegans* is by far one of the most important model organisms in many fields of research. They were the first species to ever have their complete genome sequenced [40] and they remain the only organism to have their entire cell lineage mapped, from the beginning of development within the egg to the last cell division. *C. elegans* populations can be maintained either in liquid or on agar medium with a lawn of *Escherichia coli* (*E. coli*) as its source of nutrient. Each worm undergoes the same developmental cycle, with 12 hour periods of growth between each of the four larval stages before reaching adulthood. At 48 hours post-hatch, a worm will have reached its reproductive stage in life to begin offspring production for 4-5 days [41]. Given that *C. elegans* are hermaphrodites, each worm will lay approximately 250-300 offspring that are genetic clones of the mother, barring any genetic mutations that might arise. At five days old, the reproduction window is closed, and the worm will live an additional 16-20 days on average when cultured at 20°C. The *C. elegans* lifespan is temperature dependent, where lifespan is extended when cultivated at 16°C and reduced at 25°C. In the wildtype N2 strain, only 0.1% of the population are male that arise from spontaneous X chromosome nondisjunction

events during hermaphrodite reproduction. However, mating between a male and hermaphrodite worm will result in a 50:50 ratio of male and female offspring, which presents a useful tool for introducing mutations between different genetic backgrounds.

All of these features were undiscovered until Sydney Brenner raised awareness of the usefulness of the species in 1974. His influential research displayed many different benefits of the species, including generating hundreds of different strains of *C. elegans* mutants that are utilized in different areas of research ranging from neuroscience to developmental biology [42]. Brenner introduced *C. elegans* as a model organism not only because of its simplicity, but also because of its intricate genetics. While a 1,000 cell worm may seem vastly different from humans, 83% of the *C. elegans* proteome shares homology with the human proteome [43], allowing for rapid and inexpensive research on human-homologous genes. *C. elegans* have a powerful genetic system that can be easily modified by a variety of methods: RNA interference (RNAi) for gene knockdown, chemical mutagenesis for nucleotide mutations, CRISPR/Cas9 for precision gene insertion/editing, and gene overexpression via transgenesis. Each of these genetic modifications can be stably integrated into the genome and are passed down to the offspring, allowing for genetically identical populations to be produced from a single worm. Perhaps most useful for genetic research is the transparency of this species. Tagging a fluorescent reporter to a gene results in the *in vivo* tracking of the tissue and cellular localization of the gene that can be easily visualize by stereomicroscopy. Stable strains of *C. elegans* can also be produced with reduced or overexpression of a gene that can result in impaired or improved cellular or metabolic processes, unique phenotypic differences for assay use, or a combination of multiple genetic mutations.

Aging and lifespan go hand in hand in *C. elegans*, where changes in genetics or environment can lead to an altered lifespan. Klass was the first to suggest effects of aging alteration in *C. elegans*, identifying five mutants that had an altered lifespan when compared to wild type worms [44]; Friedman and Johnson later mapped all five mutants to the same genetic locus – *age-1*, a phosphatidyl inositol 3-kinase (PI3K) homolog [45]. The research performed on *age-1* was the first to propose that individual genes could alter lifespan in *C. elegans*. Since then, a wide variety of genetic loci and pathways have been identified that regulate lifespan. One pathway of interest is the pathway used by dauer formation 16 (*daf-16*), a homolog to human

forkhead box transcription factors class O (FOXO). This transcription factor acts to elicit signals in dauer formation in *C. elegans*, as well as aging and development. Outside of the nematode, mutations in the human FOXO ortholog FoxO3A have been associated with longevity, as well as in other species such as *Drosophila* [46]. Overall, the wide variety of techniques and research available on *C. elegans* make this simple nematode a perfect model organism for aging research on a small and rapid scale.

1.4: Translation suppression and lifespan extension

It has been known for decades that lifespan extension in various organisms can be achieved by switching cellular mechanisms from a state that promotes growth to one that promotes stress resistance and repair. However, until 2005, it was thought that only caloric restriction could induce a repair-promoting state. Kaeberlein et al. were the first to challenge this idea, identifying six genes related to the target of rapamycin (TOR) pathway that increased lifespan when knocked down in yeast [47]; TOR is a kinase that regulates lifespan in many organisms through the modulation of translation, autophagy, and other cellular processes when presented with a change in nutrient levels or environment. Kaeberlein et al.'s data opened up a field of research focused on studying the effects of protein synthesis suppression on aging.

In *C. elegans*, lifespan extension has been seen through multiple methods, including dietary restriction [48], strict upregulation of stress resistance pathways [49], and through protein synthesis suppression. Hansen et al. were some of the first to recognize that a cell undergoing normal protein synthesis is a cell undergoing healthy growth, hypothesizing that the suppression of this energy-dense process may allow for reduced growth and increased cellular repair [50]. Through a series of genome-wide screens and lifespan assays with ribosomal genes knocked down, they determined that suppression of protein synthesis does indeed extend lifespan in *C. elegans* through one of two pathways, dependent on the genes knocked down [50]. In the same year, Pan et al. also found an increase in lifespan when protein synthesis is suppressed in *C. elegans* by RNAi silencing of different translation related genes (Pan et al., 2007). Since then, the topic has seen an increase in research due to the ease of protein synthesis suppression in *C. elegans* through simple means such as RNAi. While plenty of literature exists on this topic, they all share the same two common results: suppression of translation leads to an increase in lifespan

due to the reduced overall translation while also increasing translation of certain stress related genes [35].

One gene in particular presents a unique case for lifespan extension – *ifg-1* in *C. elegans* that is homologous to the human eIF4G3. This translation factor directs mature mRNA to the ribosome, and as such is the first protein to enact in the process of protein translation. When *ifg-1* is suppressed via RNAi, lifespan is greatly increased in *C. elegans* from 20 to 30 days while also delaying processes such as reproduction [51]. Knocking down *ifg-1* allows for differential expression of various genes, specifically an increase in expression of those related to stress response with long mRNA sequence lengths [52]. Past literature exists supporting the place of *ifg-1* in lifespan regulation, however it is unknown as to how this happens; the mechanism of how it regulates lifespan, if it uses the *daf-16*/FOXO pathway, or other factors that come into play. Recent research on this topic indicates that *ifg-1* suppression results in enhanced lifespan through a mechanism that heavily requires the heat shock factor *hsf-1* [53], although nothing further is known.

In this thesis, I present new research data that provides novel insights into how translational suppression acts to enhance *C. elegans* resistance to stress-induced RNA mis-splicing and describe the genetic way regulating this process.

2. Hypothesis and Objectives

2.1: Hypotheses

RNA mis-splicing and disease state is strongly hypothesized to be linked, however the mechanism behind this link is almost entirely unknown. My research aims to explore this mechanism through the environmental stress cadmium, which is known to cause RNA mis-splicing in *C. elegans*, among other effects [14]. By disrupting the RNA splicing pathway through genome-wide RNAi knockdown, I was able to investigate the effects on RNA splicing caused by knocking down individual genes, first by focusing on the entire genome and second by focusing on key genes of interest. Through this precise disruption and knockdown, I was able to identify the *ifg-1* (initiation factor 4G (eIF4G) family-1) gene, one previously known to have a prominent in lifespan regulation and extension. I was then able to further detail the mechanism through which *ifg-1* enacts its RNA splicing protection under stress through a series of assays and data analysis on RNA sequenced from *ifg-1* mutant worms.

Hypothesis 1: Genetic regulators of stress-induced RNA mis-splicing can be identified by a genome-wide RNAi screen, as well as further characterized by various assays.

Hypothesis 2: The *ifg-1* gene in *C. elegans* has a role in RNA splicing protection under stress, which can be further elucidated.

2.2: Objective 1

Identify and characterize all genes in the *C. elegans* genome that, when knocked down, improve the organism's RNA splicing fidelity under cadmium-induced stress. As well, choose a subset of the identified genes, determined by previous literature available, to characterize their effect on lifespan under normal and stressed (cadmium) conditions when knocked down.

2.3: Objective 2

Characterize the mechanism of lifespan regulation and extension provided by *ifg-1* in *C. elegans*, understand how *ifg-1* affects RNA splicing under stress, and define the role of *ifg-1* in RNA splicing protection.

3. Materials and Methods

3.1: Worm growth conditions

All strains of *C. elegans* were grown on nematode growth media (NGM), adapted from [56]. Strains were obtained from the *Caenorhabditis* Genetics Centre (CGC) or self-made (Table 3.1). Worms were fed interchangeably with a slow-growing lab strain of *E. coli* (OP50) or a faster growing strain of *E. coli* (NA22) (Table 3.2) unless otherwise stated. Worms were transferred to fresh *E. coli* plate when their plate was spent (either by depletion of food or by damage to the agar) via different methods depending on the downstream application. For transferring between experimental plates, worms were transferred via washing in M9 buffer (Table 3.3) or water. For lifespan assays, worms were picked individually with a platinum wire pick or a hair pick [56]. For general maintenance, a piece of agar with starved worms was excised with a sterile tool and placed on a fresh plate. Worms were kept at 20˚C unless otherwise stated. Worms were observed using the Olympus SZX2-ILLT (Olympus) microscope unless otherwise stated.

3.2: Bleach Synchronization

To obtain a population of identical worms all at the L1 stage of life, bleach synchronization [56] is performed. This process dissolves the adult worms to release fertilized eggs which are protected from the lysis solution by its egg shell. A large population of adult worms each retaining ~10 eggs each was washed off the plate into sterile 15 mL polyproline conical tubes (Falcon) using sterile water. Tubes were then spun at 3500 *x g* for 15 seconds to pellet worms at the bottom of the tube. Supernatant was removed and the worm pellet was resuspended with fresh sterile water to repeat the wash step. This washing step was repeated until the supernatant was clear, indicating the removal of bacteria. On the final wash, water was removed until there was 5 mL total volume left in the tube. At this point, a worm lysis solution of 0.3 mL of 1M NaOH and 1.3 mL of 100% bleach were added to the tube before being shaken vigorously for approximately five minutes. After adults had dissolved (maximum eight minutes), filter sterilized water was added to fill the tube, then spun down for two minutes at 3500 *x g* to neutralize the lysis solution. Supernatant was then immediately poured off, and sterile water was

Table 3.1: Strains used in this study. Strain names, genotypes, purpose, and where they were obtained for all worm strains used in this research.

Strain Name	**Genotype**	**Purpose**	**Source**
N2 Bristol	Wild-type	Control strain	*Caenorhabditis* Genetics Centre (CGC)
KH2235	*lin-15 (n765) ybIs2167* [eft-3::RET-1E4E5(+1)E6-GGS6-mCherry eft-3::RET-1E4E5(+1)E6-(+2)GGS6-EGFP lin-15 (+) pRG5271Neo] X	Fluorescent strain used as a biomarker for RNA splicing errors	Kuroyanagi lab
KX54	*ifg-1(cxTi9279)* II; *bcIs39 bcIs39 [lim-7p::ced-1::GFP and lin-15(+)] V* V.	Used to create the *ifg-1* strain	Keiper lab
ifg-1	*ifg-1(cxTi9279)*	Investigating the effects of *ifg-1* knock-down	Wu lab

Table 3.2: Reagents used in this study. Reagent names and where they were obtained for all reagents used in this study.

Reagent	**Source**	**Catalog Number**
6X DNA loading dye	Thermo Fisher Scientific	FERR0611
96-well plates	Falcon	08-772-2C
Agar	Thermo Fisher Scientific	DF0479-17-3
Agarose	Thermo Fisher Scientific	BP160-500
Bactone peptone	Thermo Fisher Scientific	DF0118-17-0
β-mercaptoethanol	Thermo Fisher Scientific	PI35602
Cadmium chloride	Thermo Fisher Scientific	AC315270050
Calcium chloride	Thermo Fisher Scientific	C79-3
Carbenicillin	Thermo Fisher Scientific	BP26485
Cholesterol	Thermo Fisher Scientific	MP21013803
Conical centrifuge tubes	Falcon	Varies
Deepwell 96-well plates	Thermo Fisher Scientific	12-566-120
DNAse enzyme,	Thermo Fisher Scientific	FEREN0521
DreamTaq	Thermo Fisher Scientific	FEREP0702
EDTA	Thermo Fisher Scientific	AC327345000
Ethanol	Greenfield Global	N/A
Glycerol	Thermo Fisher Scientific	PI17904
Hydrochloric acid	Thermo Fisher Scientific	SA48-500
Individually sealed pipettors	Thermo Fisher Scientific	Varies
IPTG	Thermo Fisher Scientific	15529019
Lysogeny broth	Thermo Fisher Scientific	DF0446-17-3
Magnesium chloride	Thermo Fisher Scientific	M33-500
Magnesium sulphate	Thermo Fisher Scientific	M65-3
Microcentrifuge tube	Corning	Varies
Microscope cover slips	Thermo Fisher Scientific	12-545-87
Microscope slides	Thermo Fisher Scientific	12-550-A3
MultiScribe reverse transcriptase	Thermo Fisher Scientific	4311235
NA22 bacteria	CGC	N/A
NP-40	Alfa Aesar	AAJ60766AK

OP50 bacteria	CGC	N/A
PCR grade water	Thermo Fisher Scientific	10977015
Petri dishes	Corning	Varies
Pipette tips	Thermo Fisher Scientific	Varies
Platinum wire, 0.25mm	Alfa Aesar	AA45093BY
Potassium chloride	Thermo Fisher Scientific	P217-3
Potassium phosphate	Thermo Fisher Scientific	P285-3
PowerUp SYBR green master mix	Thermo Fisher Scientific	A25778
Proteinase K	Thermo Fisher Scientific	25530031
qPCR primers	Eurofin Genomics	Varies
RNA purelink micro kit	Thermo Fisher Scientific	K310250
RNAi bacteria library	[54], [55]	N/A
Sodium azide	Thermo Fisher Scientific	BP922I-500
Sodium chloride	Thermo Fisher Scientific	18606413
Sodium phospate	Thermo Fisher Scientific	S374-1
SYBR I nucleic acid gel stain	Thermo Fisher Scientific	S7563
TAE	Thermo Fisher Scientific	FERB49
Tris	Thermo Fisher Scientific	BP152-500
Tween-20	Thermo Fisher Scientific	AAJ20605AP

Table 3.3: Medias and their recipes used in this study. Reagents were added from top to bottom, stirred between each addition, and autoclaved before adding reagents listed after "autoclave".

Media	**Recipe source**	**Recipe**
M9 Buffer	[56]	1 L H_2O 5.0 g NaCl 3.0 g KH_2PO_4 6.0 g Na_2HPO_4 Autoclave 1 mL $MgSO_4$ (1M)
NGM Buffer	[56]	1 L H20 3.0 g NaCl Autoclave 1 mL Cholesterol (1M) 1 mL $CaCl_2$ (1M) 1 mL $MgSO_4$ (1M) 25 mL KH_2PO_4 (1M, pH 6)
NGM Plates	[56]	1 L H20 3 g NaCl 2.5 g Peptone 25 g Agar Autoclave 1 mL Cholesterol (1M) 1 mL $CaCl_2$ (1M) 1 mL $MgSO_4$ (1M) 25 mL KH_2PO_4 (1M, pH 6)
NA22 Plates	[56]	1 L H2O 1.2 g NaCl 20 g Peptone 25 g Agar Autoclave

		1 mL Cholesterol (1M) 1 mL $CaCl_2$ (1M) 1 mL $MgSO_4$ (1M) 25 mL KH_2PO_4 (1M, pH 6)
Lysis Buffer	[57]	3.73 g KCl 0.508 g $MgCl_2$ 10 mL 1M Tris-HCl, pH 8.8 4.5 mL NP-40 4.5 mL Tween-20
100x SYBR Green	Thermo Fisher	2 uL 10,000x SYBR Green Dye 198 uL 1x TAE buffer
SYBR Green + Loading Dye	Thermo Fisher	100 uL 100x SYBR Green 200 uL 6x Loading Dye
DNA Ladder	Thermo Fisher	30 uL SYBR Green + Loading Dye 90 uL H_2O 10 uL DNA Ladder
Primer Mix	Eurofin Genomics	10 uL forward primer (100 µM) 10 uL reverse primer (100 µM) 180 uL RNAse free water

added up to 10 mL. The tube was flicked to resuspend the pellet of eggs to ensure thorough washing of each egg, then centrifuged again for two minutes at 3500 *x g*. This wash step was then repeated once more to ensure no bleach remained in the tube. Supernatant was then removed until 500 µL remained, then the last 500 µL was mixed with the egg pellet and deposited onto a 6 cm unseeded NGM plate to allow eggs to hatch overnight. After 18 hours, all viable eggs would have hatched and are synchronized at the L1 stage, which can be washed off the plate with sterile M9 for experimental use.

3.3: RNAi screen

To perform a large-scale liquid RNAi screen [60], the KH2235 *C. elegans* strain were grown under normal conditions and bleach synchronized to obtain a large L1 population. RNAi *E. coli* was grown in 600 µL Lysogeny Broth (LB) containing 50 mg/mL carbenicillin for 18 hours at 37˚C in 96-well deep well plates. After this, *E. coli* were mixed with 25 µL 100 mM Isopropyl β- d-1-thiogalactopyranoside (IPTG) was added to each 96 well containing the RNAi bacteria and incubated for one hour at 37˚C to induce the dsRNA plasmids within the bacteria. Bacterial plates were then spun down at 1572 *x g* for 20 minutes to pellet bacteria, supernatant was removed, and bacteria was resuspended in 100 µL NGM buffer containing 4 mM IPTG and 0.05 mg/mL carbenicillin. Thirty µL of this bacterial solution was then transferred into a new 96 well microplate, and 70 µL of the same NGM solution containing IPTG and carbenicillin with approximately 30 L1 KH2235 worms suspended in it was added to each well. Plates were then placed on an orbital shaker in 20˚C for 48 hours to allow for worm growth and exposure to RNAi. After 48 hours, each well was screened under an Olympus SZX2-ILLB fluorescent microscope for changes in GFP and RFP. Wells without visible GFP fluorescence were noted at this stage to determine RNAi that suppress basal RNA splicing. Three hundred µM cadmium chloride solution was then added to each well to act as a chemical stress to inhibit RNA splicing, and returned to the previously described conditions for an additional 24 hours to allow for cadmium exposure. After 24 hours, each well was screened again under the same fluorescent microscope, however this time populations were screened for their ability to retain GFP fluorescence. Those that retained fluorescence were marked down, as well as the intensity of their fluorescence on a three-point scale (strong, medium, and weak) and the size of the worms

on a three-point scale (large, medium, small). Wells were then screened for a loss of RFP, and wells without RFP were marked down.

To eliminate no false-positive results from the primary RNAi screen, a secondary RNAi screen containing only the RNAi hits identified in the primary screen was performed, however this time occurring on NGM agar plates instead of in liquid NGM. Approximately 50 L1 synchronized KH2235 worms were placed on a 6 cm NGM plate containing 1 mM IPTG and 0.05 mg/mL carbenicillin, seeded with the corresponding RNAi *E. coli*. After 48 hours of growth at standard conditions, worms were moved to a corresponding RNAi *E. coli* seeded 6 cm NGM plate containing 300 μM cadmium chloride. After 24 hours of cadmium exposure, worms were screened for their GFP and RFP fluorescence and scored on the same three-point scale as the initial RNAi screen. Gene knockdowns that enabled a population of worms to retain fluorescence were considered knockdowns of interest, and a glycerol stock of the original RNAi *E. coli* was taken for further use by mixing 10 μL of grown bacteria with 1 mL of LB containing 1 mM IPTG and 0.05 mg/mL carbenicillin and 50% glycerol and preserved at -80˚C.

3.4: Genetic Crosses

While *C. elegans* are typically hermaphrodites, a male can spontaneously hatch in a population due to a rare spontaneous loss of the X-chromosome (~0.1%). One way of sharing genetic information between strains is to mate a male worm with a hermaphrodite, and select for offspring that have retained the desired combination of genetic information. However, one male per every thousand hermaphrodites results in an extremely inefficient method of genetic crosses; instead, males in a population can be enriched through heat shocking L4 hermaphrodites to induce spontaneous loss of the X chromosome in its offspring. To enrich for males, approximately 40 worms were placed onto a 10 cm NGM plate seeded with OP50 bacteria and heat shocked at 30˚C for 6 hours, then move to a 20˚C incubator for maintenance. This heat shock allows for the hermaphrodites to produce and lay eggs containing males. After two days, plates are checked daily for male offspring. If found, the male offspring is moved to a 6 cm plate with 100 μL bacteria seeded directly in the centre of the plate (a “mating plate”), then one L4 hermaphrodite from the same strain for every two males is added. Three more plates were

prepared, each containing a total of three hermaphrodites and six males. This “self-cross” was repeated every three to four days to maintain a continuous healthy population of males.

The KX54 worms contain homozygous *ifg-1* partial loss of function mutation, but also carry an unwanted GFP marker used by another lab. To isolate the *ifg-1*(cxTi9279) genotype, six adult N2 males and three L4 KX54 hermaphrodites were then placed on a 6 cm mating plate, and repeated three times for four replicates total. After allowing worms to mate for 48 hours, nine F1 offspring were removed from one of the four plates and placed onto individual plates. If mating took place, the F1 generation will be heterozygous for the *ifg-1* mutation and the GFP marker, which will segregate in a Mendelian fashion that will permit the isolation of F2 carrying only the *ifg-1* mutation. A PCR method was then used to genotype the *ifg*-1 mutation. To do this, several single F2 worms were individually placed onto an agar plate and allowed to lay F3 offspring, after which the F2 worms were lysed to perform single worm genotype PCR. The F2 worm was frozen in 10 μL of lysis buffer (Table 3.3) with 1% Proteinase K added before use for ten minutes at -80˚C, then incubated for one hour at 65˚C to initiate worm lysis and for an additional 15 minutes at 95˚C for proteinase K denaturation. After this, gel-based PCR was performed using primers designed for the genetic mutation of interest (Table 3.4) to determine if the correct mutation was retained in the F2 hermaphrodite. The F3 offspring that are genetically identical to the single F2 worm that exhibited homozygous mutation to *ifg-1* was isolated for downstream experiments.

3.5: Lifespan Assays

Lifespan assays were adapted from procedures detailed in [61]. A population of synchronized N2 worms was grown on 10 cm NGM plates containing 1 mM IPTG and 0.05 mg/mL carbenicillin, and seeded with 1 to 2 mL empty vector (EV) bacteria. EV bacteria expresses a plasmid that is not homologous to any *C. elegans* genes and serves as the control *E. coli* for RNAi experiments. After 48 hours of growth, approximately 40 worms were moved to 6 cm plates seeded with 500 μL of the *E. coli* RNAi of choice (Day 1). The populations were then checked each day for the following parameters: offspring presence, *E. coli* remaining, or dead worms. If offspring were present on the plate or if the plate was showing signs of bacteria depletion (visible first on the edges of the plate), the original ~40 adult worms were moved to a

Table 3.4: Primers used in this study. Primers were all sourced from previous research – the validity and efficacy of each primer can be found in the source papers referenced.

Gene of interest	**Source**	**Primer sequence**
ret-1	[58]	F: 5'-CATCCGCTGAAGGATCCATAG-3' R: 5'-GAGCTTCCTCAGCAATCGGAG-3'
ifg-1 genotype	[59]	F: 5'-ACCAAACTGGGCAAACAAAG-3' R: 5'-CTTCCTGAAATTTGGTTTAACAGT-3'
oJL 115 mos	[59]	R: 5'-GCTCAATTCGCGCCAAACTTATG-3'
rpl-2 (qPCR)	[14]	F: 5'-CTTTCCGCGACCCATACAA-3' R: 5'-CACGATGTTTCCGATTTGGAT-3'
cdr-1 (qPCR)	[14]	F: 5'-ATTACTGCTGCGCTGTTTGG-3' R: 5'-GGGGACAAGTTCGGACAGTT-3'
numr-1 (qPCR)	[14]	F: 5'-AGACGTCACTGTTTTGGTGGA-3' R: 5'-CCGAATCCTCCAGTTGGACC-3'
hsp-16.49 (qPCR)	[14]	F: 5'-TGAGTTGTGATCAGCATTTCTC-3' R: 5'-GGATGAAATCACTGGATCTGTT-3'

fresh identical plate via picking with a flame-sterilized platinum pick. After approximately six days (for N2 worms, this age varies for each gene knockdown), the worms should not lay any more offspring and can instead be monitored for deaths. If a worm is suspected to be dead, they can be checked for life by one of two ways: either by gently tapping the plate on the microscope stage to induce vibration or by gently poking the tail region with a sterilized platinum pick. If no response is seen after both actions, the worm is considered dead and the age of death in days is marked down.

In some cases, worms may die due to unnatural causes such as human errors, crawling up on the side of the plate where there is no agar and drying out, or developmental abnormalities that result in vulva defects leading to protrusions or bagging phenotypes. When a worm dies due to any unexpected cause, it was removed from the plate and was marked as censor instead of death. Counting the number of censor worms helps to accurately trace the lives of all worms in a population, and plates with an extreme number of censors were discarded due to the statistical error introduced by too high a number of censored worms. Lifespan data were analyzed by GraphPad Prism version 8.2.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com to generate lifespan curves and by the Online Application for Survival Analysis 2 (OASIS 2) [62] for Kaplan-Meier estimator statistics.

3.6: Cadmium Sensitivity Assays

A second endpoint used in my studies was the cadmium survival of a population of worms, where survival is determined by the age of a population under cadmium conditions. Cadmium sensitivity assays were performed nearly identically to lifespan assays [61], with one key difference where worms are moved to cadmium containing agar plates after reaching adulthood. N2 worms were bleach synchronized, allowed to grow for 48 hours on 10 cm NGM plates containing 1 mM IPTG and 0.05 mg/mL carbenicillin seeded with EV bacteria, then moved to similar 6 cm plates seeded with the appropriate RNAi bacteria, just as in lifespan assays. However, after 48 hours on RNAi knockdown conditions, worms were moved once again to 6 cm plates containing 300 μM cadmium chloride and seeded with 500 μL of the appropriate RNAi *E. coli*. After this transfer, cadmium sensitivity assays are carried out identically to

lifespan assays, however they are much shorter in duration and require no transfer to new plates due to toxicity caused by cadmium exposure.

Further on in this research, cadmium sensitivity assays were performed that included the slow-growing *ifg-1* mutant strain (Table 3.1). To account for the slower rate of development, *ifg-1* worms were allowed to grow for 72 hours on EV bacteria seeded plates before moving to 6 cm RNAi seeded plates; if left to grow for only 48 hours as N2 worms were, they would be moved to cadmium conditions while they are too young. Exposure to cadmium at a young age can result in the halting of growth and death at a younger age in *C. elegans*, so an extra 24 hours growth on EV food was necessary to avoid these issues. Cadmium sensitivity data was analyzed in the same manner as described above using GraphPad and OASIS2.

3.7: RNA Extraction

N2 and *ifg-1* mutant worms were bleach synchronized and allowed to grow for 48 hours on EV bacteria. After 48 hours, the populations were moved to plates that either did or did not contain 300 μM cadmium, seeded with the same bacteria (total of four replicates per condition). After 24 hours of cadmium exposure, worms were then washed off their plates using M9 containing 1% LB, which prevented the worms from sticking to the plastic tubes or pipette tips. Worms were rinsed three times in M9 with 1% LB, then transferred to 1.5 mL microcentrifuge tubes (approximately 2,000 worms per tube) and all supernatant was removed. After this, 300 μL of ice-chilled lysis buffer from the Invitrogen™ PureLink™ RNA Mini Kit containing 1% β-mercaptoethanol was added to each sample and immediately sonicated for 15 seconds twice, with a brief incubation period on ice between homogenizations to prevent heat denaturation of RNA. Samples were checked under the microscope after the second homogenization to ensure full degradation of worms. The sonicator and area was cleaned with RNAse away solution and 70% ethanol between conditions to ensure no cross-contamination of samples.

RNA was then extracted from the homogenized samples via the Invitrogen™ PureLink™ RNA Mini Kit. To determine the quality of the extracted RNA, samples were separated and visualized on a 2% agarose gel (120 volts, 30 minutes) using SYBR green I nucleic acid gel stain (Thermo Fisher). Wells that contained two distinct bands at 5,000 bp and 2,000 bp

(corresponding to 28S and 18S, the two constituents of the ribosome) were considered successes, and wells that did not were considered failures and RNA extraction was performed again.

DNAse treatment was then performed using DNAse I, RNAse Free (Thermo Fischer Scientific). A maximum of 1 μg normalized RNA was added to a 200 μL PCR tube (up to 8 μL), as well as: 1 μL buffer; 1 μL DNAse enzyme; and RNAse free water up to 10 μL. Tubes were mixed thoroughly and centrifuged briefly, then incubated for 30 minutes at 37˚C in the ProFlex PCR System (Applied Biosystems). At this point, 1 μL of 50mM ethylenediaminetetraacetic acid (EDTA) was added, and samples were then incubated for an additional 10 minutes at 65˚C. At this point, DNAse treatment had been completed, and cDNA conversion was immediately performed.

cDNA conversion was performed using MultiScribe™ reverse transcriptase (Applied Biosystems). To each reaction tube, 5 μL of DNAse treated RNA was added, as well as the following: 1 μL buffer; 0.4 μL dNTP mix; 1 μL random primers; 0.5 μL MultiScribe reverse transcriptase; and 2.1 μL RNAse free water. Samples were centrifuged briefly and spun down, loaded into a ProFlex PCR System (Applied Biosystems), and incubated at the following settings: 10 minutes at 25˚C; 120 minutes at 37˚C; and 5 minutes at 85˚C. After incubation was complete, 90 μL of RNAse free water was added to each tube. After cDNA was synthesized, gel-based PCR was performed to determine the splicing pattern of *ret-1*.

3.8: Gel-Based PCR

Gel-based PCR was used as a method of validating endogenous RNA splicing or to genotype worm mutations. In each 200 μL PCR tube, the following was added: 1 μL sample DNA; 0.2 μL dNTP; 1 μL primer mix (10 μL forward primer, 10 μL reverse primer, 180 μL RNAse free water); 0.05 μL Dream Taq polymerase; and 6.75 μL RNAse free water. Tubes were thoroughly vortexed then centrifuged briefly before being loaded into a ProFlex PCR System and incubated with the following PCR method: 1 minute at 95˚C; 30 seconds at 95˚C; 30 seconds at 60˚C; 1 minute at 72˚C; 10 minutes at 72˚C. Generally, 30-35 cycle settings were used to amplify the PCR product. Once complete, 3 μL of 6X DNA loading dye (Thermo Fisher Scientific) spiked with SYBR nucleic acid I gel stain (Thermo Fisher Scientific) was added and all 13 μL of sample was separated on a 2% agarose gel for 45 minutes at 100V using a PowerPac

Basic electrophoresis chamber (Bio-Rad). Gels were imaged using a ChemiDoc MP Imaging System (Bio-Rad).

3.9: qPCR

qPCR was performed using PowerUp™ SYBR™ Green (Thermo Fisher Scientific) to confirm that cadmium exposure experiments induced toxicity by measuring the expression of known cadmium-inducible genes: *cdr-1, numr-1, hsp16.49,* and *rpl-2* was used as a housekeeping gene. In each well in use in a MicroAmp Fast Optical 96-Well Reaction Plate with Barcode (0.1mL) (Applied Biosystems), the following was added: 2 µL sample cDNA; 5 µL PowerUp™ SYBR™ Green; 1 µL Primer mix (10 µL forward primer, 10 µL reverse primer, 180 µL RNAse free water); and 2 µL RNAse free water. Plates were then sealed with an optical adhesive cover (Applied Biosystems) ensuring a complete tight seal around each well, before tapping gently against the countertop and centrifuging briefly. The microplate was then placed into a QuantStudio 3 qPCR machine (Applied Biosystems), where a Comparative CT experiment was set up with the chemistry type of SYBR green reagents. The following method was then ran with the correct plate layout and ROX as the passive reference: hold stage: 50˚C for 2 minutes, 95˚C for 10 minutes; PCR stage: 95˚C for 15 seconds, 60˚C for 1 minute repeated 40 times; melt curve stage: 95˚C for 15 seconds, 60˚C for 1 minute, and 95˚C for 15 seconds.

3.10: RNA-sequencing

RNA was extracted from the following conditions and sent of RNA-sequencing analysis to characterize their transcriptomes: 1) N2 + control, 2) N2 + cadmium, 3) *ifg-1* + control, and 4) *ifg-1* + cadmium. Three replicate RNA samples for each experimental condition were shipped on dry ice to Novogene where the sequencing was experiments were performed. When RNA was received at Novogene, a poly-A enriched cDNA library was constructed from the total RNA followed by next generation sequencing. Once sequencing was complete, bioinformatics analysis took place. Between each step, a unique quality control check took place to ensure high quality samples and results. Genes that were differentially expressed and alternatively spliced were analyzed and compared between the four conditions listed. More detail about the steps of RNA-seq can be seen in Appendix A.

3.11: *C. elegans* imaging

To image *C. elegans,* two methods can take place. First, animals were imaged while on their source plate by placing the plate under an Olympus SZX2-ILLB fluorescent microscope and imaged using the QImaging Retiga R3 (Cairn Research Ltd) camera. To obtain higher quality images of multiple aligned worms, worm mounting was required. To do this, a drop of liquid 2% agarose was placed on a microscope slide and another microscope slide was immediately placed on top to press the drop and create a flat surface. After a brief cooling period (~2 minutes), the top microscope slide was removed and 10 μL of 2% sodium azide dissolved water was placed on the agarose platform to anesthetize the worms. Generally, 8-10 animals were mounted on a single slide for imaging with a QImaging Retiga R3 (Cairn Research Ltd) attached to a ZEISS Axio Vert.A1 Microscope (Zeiss).

3.12: Statistical analysis

Strength of fluorescence and size of worm was determined qualitatively for RNAi screens. Functions of genes were clustered by The Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Database (Version 6.7) [65], [66]. For lifespan and cadmium sensitivity assays, survival curves were generated via GraphPad Prism version 8.2.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com and quantitative data was determined by log-rank tests via the OASIS online tool [62]. Changes in lifespan or survival were compared to each other qualitatively via the average lifespan or survival to determine differences in lifespan. Student's t-test was used to compare one independent variable; however, one-way ANOVA was used to compare qPCR data. Two-way ANOVA was used to compare two independent variables. The Student's t-test and ANOVAs were all performed via GraphPad Prism version 8.2.1 for Windows. Pearson's correlations were determined via GraphPad Prism version 8.2.1 for Windows. All experiments were performed in triplicate unless otherwise stated. Across all experiments, * indicates $P < 0.01$, ** indicates $P < 0.001$, *** indicates $P < 0.0001$, and **** indicates $P < 0.00001$.

4. Results

4.1: RNA splicing is protected via translation suppression

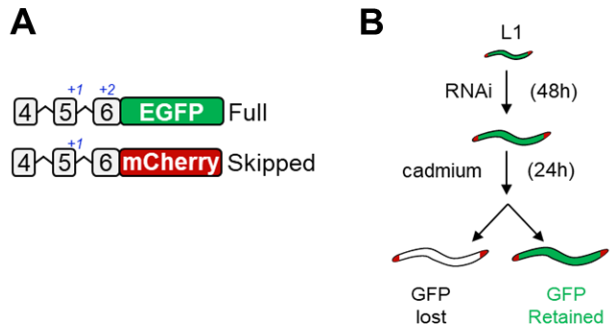
In order to determine the effect of single-gene knock-down on the RNA splicing mechanism in *C. elegans*, a genome-wide RNA interference (RNAi) screen was performed (adapted from [60]). This is accomplished by using an *in vivo* RNA splicing strain (KH2235, Materials and Methods Table 3.1) which expresses a dual fluorescent reporter tagged to two spliced variants of a *C. elegans* gene called *ret-1*. In brief, this splicing reporter can be used to detect exon skipping that results in the decrease in GFP fluorescence but increase in mCherry fluorescence (Figure 4.1A). This is accomplished by +1 or +2 nucleotide insertions to exon 5 and 6 of the *ret-1* gene that dictates the downstream fluorescent protein reading frame. It was previously shown that exposure of this splicing reporter to the heavy metal cadmium completely suppress GFP fluorescence and inhibits RNA splicing [14], indicating the usefulness of cadmium as a compound to study mechanisms of stress-induced RNA splicing inhibition.

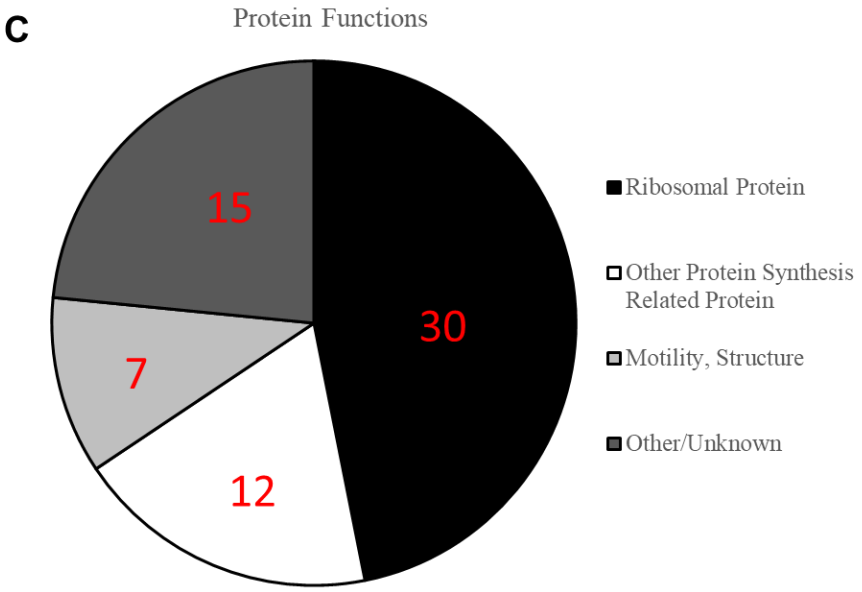
The workflow of the RNAi screen is shown in Figure 4.1B, where a simplified overview is depicted. The goal of the screen is to identify genes that when knocked down by RNAi can provide resistance against cadmium-induced RNA splicing disruption. A total of 137 hits were recovered by the primary RNAi screen, this number was narrowed down to a total of 64 unique genes after a secondary screen of the 137 initial hits in triplicates. The discrepancy in this number can be attributed to a few factors: repeated identical genes that were present in multiple wells in the RNAi screen as a form of quality control (4 wells); wells of bacterial stock that did not have an annotated gene associated with them or the gene was unknown (13 wells); wells of bacterial stock that would not re-grow a population of bacteria in the secondary screen (1 well); or genes that did not reproduce fluorescence when knocked down in the secondary screen, thus being deemed as false positives (55 wells). Only the 64 genes confirmed by the secondary RNAi screen were used for further research due to the unknown or unstable outcomes produced by the other 73 wells. Out of the 64 genes identified, a stunning 42 of them are genes with known functions related to protein synthesis, and 30 of the genes encode for constituents of the ribosome. The remaining 22 genes related to protein synthesis include those functioning as tRNA synthetases, elongation factors, and initiation factors. These data are summarized in Figure 4.1C

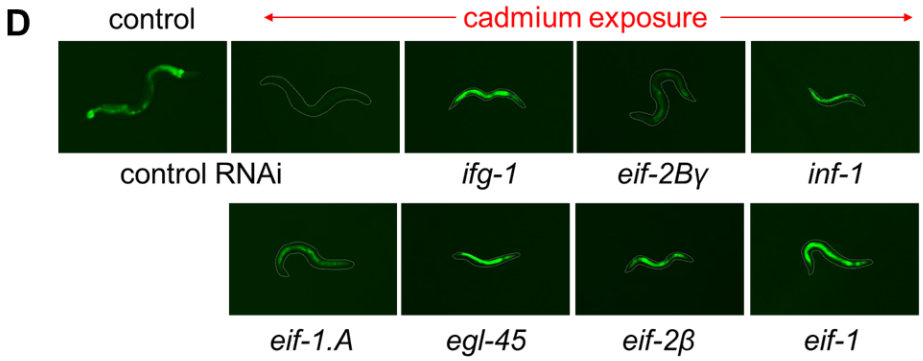
and Figure 4.1E, where the functions of the genes of interest are categorized, or in Appendix B where the functions and ontology of these genes are described in greater detail.

Data was also collected on the relative strength of fluorescence and on the developmental stage of *C. elegans* after knockdown and cadmium exposure for the 64 genes identified; these data are summarized in Appendix B. The 64 genes of interest were further categorized for their relative effect on protection against cadmium-induced decrease in GFP signal. In populations where the gene knock-down caused a slight protection of RNA splicing, a faint GFP response was present; whereas in populations were the gene knock-down caused near full protection of RNA splicing fidelity, a strong GFP signal similar to pre-cadmium exposure was present. Out of the 42 genes related to protein synthesis identified here, 25 produced a strong GFP fluorescence, 12 produced a medium strength fluorescence, and 5 produced a faint fluorescence. For the 22 genes unrelated to protein synthesis, 11 produced a strong GFP fluorescence, four produced a medium strength fluorescence, and six produced a faint fluorescence.

The 64 genes identified were also queried for their effect on the growth and reproduction effect after the 48 hour cadmium exposure in the RNAi screen. In *C. elegans*, worm size is in direct relation to the stage of life of the worm. As such, assessing the size of the worm at a certain point in life can indicate how well the worm is able to grow and develop under the conditions specified; as a baseline, N2 wild-type worms take approximately 48 hours to reach egg-laying adulthood after hatching. Although the worms had 48 hours to grow before cadmium exposure, none of the 64 gene knockdowns of interest allowed the populations to reach egg-laying within the 48 hours given, and only six of the recorded 56 sizes were near adulthood. The rest of the recorded populations remained arrested in the L1 to L3 stage of life. This is not surprising given that majority of the 64 genes are considered to be involved in essential cellular processes (i.e. protein translation), and knockdown of these genes prohibited the proper development to adulthood. Overall, this RNAi screen revealed a large number of genes with functions related to protein synthesis, when knocked down via RNAi, lead to the protection of RNA splicing fidelity as evident by using an *in vivo* GFP reporter.







E

Functional annotation	Term	Fold enrichment	FDR
Biological process	Translation	20.3 (n=46)	6.70E-49
Cellular compartment	Cytosolic ribosome	34.5 (n=28)	3.00E-36
Metabolic function	Ribosome structure	25.2(n=29)	5.70E-31
KEGG pathway	Ribosome	13.2 (n=29)	3.90E-26

Figure 4.1: Genome-wide RNAi screen results

Results of the 64 gene hits identified in the genome-wide RNAi screen. **A)** Simplified layout of the *ret-1* reporter in the KH2235 strain of *C. elegans* depicting the construction of the two splicing transgenes. Single base pair additions to exon 5 and 6 are notated in blue. **B)** Workflow for the RNAi screen experiment, where L1 worms are allowed to grow to L4 for 48 hours before exposure to cadmium for 48 hours, when they are then scored for GFP fluorescence. **C)** Pie chart of the functions of the 64 genes identified as genes of interest in the genome-wide RNAi screen. Ribosomal proteins includes proteins making up the structure of the large and small subunits of the *C. elegans* ribosome; other protein synthesis proteins includes splicing factors, tRNA synthetases, and other essential proteins for protein synthesis; motility, structure proteins include those that make up the structure or aid in motility of the worm; and other/unknown proteins encompasses all other proteins. Functions clustered by The Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Database (Version 6.7) [65], [66]. **D)** Representative images of seven worms from the RNAi screen from populations identified as "able to retain fluorescence under cadmium" compared to control worms with cadmium exposure. *ifg-1*: Initiation Factor 4G (eIF4G); *eif-2Bγ:* Eukaryotic Initiation Factor 2Bγ; *inf-1*: Initiation Factor 1; *eif-1.A:* Eukaryotic Initiation Factor 1.A; *egl-45:* Egg Laying Defective 45; *eif-2β:* Eukaryotic Initiation Factor 2β; *eif-1*: Eukaryotic Initiation Factor 1. **E)** Major functional annotations of the 64 gene hits identified as determined by The Database for Annotation, Visualization and Integrated Discovery (DAVID) Bioinformatics Database (Version 6.7) [65], [66].

4.2: Lifespan and stress resistance is enhanced via translation suppression

An advantage of the *C. elegans* model is ease of assaying effect of a gene on aging, given their relatively short lifespan of ~ 3 weeks. Here, I used lifespan assays as a method to study the effect specific single gene knockdowns had on aging. Since *C. elegans* are hermaphrodites, it is essential to separate the aging worms from their offspring. In lifespan assays, aging worms are checked and moved to a new agar plate every day for the first 8-10 days to ensure offspring do not grow and become mixed in with the population of interest. After 10 days where self-reproduction is ceased, populations are checked every one to two days to monitor natural worm death. The final age of each worm in a population is then used to calculate the mean lifespan of the population, as well as lifespan curves that help visualize the effect on lifespan. Similarly, by adding cadmium to the agar plate, the stress resistance of a population of worms after gene knockdown can also be studied. In cadmium resistance assays, worms are first allowed to develop to adulthood under normal conditions and fed with RNAi before moving to identical RNAi plates containing a sub-lethal dose of cadmium for the rest of their lifespan.

A subset of twelve genes was chosen from the list of 64 based on their relation to translation, previous literature available, and potential for novelty; these genes and their functions are listed in the Figure 4.2 legend and described below. The gene list is split into two subsections: the first set of six genes having a role in translation and are well researched in previous literature; and the second set of six genes is less researched and the presently known functions are not directly related to translation, although more may come to light about the function of these genes in the future.

Out of the six well-researched genes that are all related to translation, *ifg-1* is one of the most studied gene on this list and is highly regarded for its role in lifespan extension and regulation. The *C. elegans ifg-1* gene is an ortholog of the human eukaryotic translation initiation factor 4 gamma 3 (EIF4G3). While predicted to have mRNA binding activity, it is known to act as a translation initiation factor and is necessary for the proper function of translation. Next, *rpl-7A* is a constituent of the ribosome, specifically the large subunit. It is an ortholog of the human ribosomal protein 7A, and in *C. elegans* it displays calmodulin binding activity, a protein with a variety of roles from muscle contraction to memory. Third is *rps-6*, a constituent of the small subunit of the ribosome, and has been researched to have a role in determining adult lifespan of

C. elegans. Next is *BMS1*, an ortholog of the human ribosome biogenesis factor *BMS1*, and is predicted to have a role in ribosome production. *rps-23* is another constituent of the small subunit of the *C. elegans* ribosome, indicating its role in translation. And lastly, F57B9.3 (gene IDs are used where gene names are not yet assigned) is an ortholog of the human eukaryotic translation initiation factor 4A2 (EIF4A2).

The second set of six genes chosen for this research have a variety of functions, most unrelated to translation. As well, they have very little literature done on them beyond genomic position mapping and basic predicted functions. Y39E4B.1 displays ATP binding activity, as well as H12I13.2, which also may have a hand in protein kinase activity. F47B3.6 shows activity with protein tyrosine phosphatase, and *nhr-61* may have a role in translation with DNA binding transcription factor activity. F46F11.1 acts as a kinase, and lastly *csp-2* has caspase binding activity. By choosing a variety of gene functions besides translation, I hoped to continue the "unbiased" efforts of this research in uncovering the link between aging and translation.

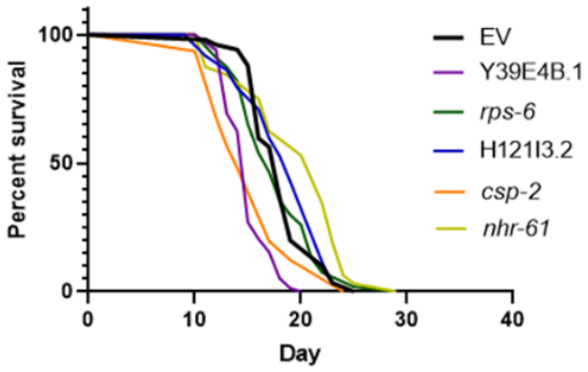
The lifespan assays performed here, as seen graphically via lifespan curves in Figure 4.2 and summarized numerically via Kaplan-Meier estimation in Table 4.1, indicate that translation suppression has a significant influence on increasing lifespan. On average, the knockdown of genes associated with translation increased the lifespan of the worm relative to the EV (empty vector RNAi) control population, with *ifg-1* or *BMS1* knockdowns showing the most pronounced increase in lifespan. Alternatively, knockdown of the six genes with functions unrelated to translation resulted in either a minor change to a decrease in lifespan. While there were some cases when knockdown of a gene from this second category resulted in a small increase to lifespan, overall the suppression of translation had a much stronger positive effect on lifespan than the knockdown of genes unrelated to translation.

To determine the effects of candidate gene knock down have on stress resistance, similar lifespan assays were performed with worm populations on agar plates containing cadmium. The results of the cadmium sensitivity assays can be seen in Figures 4.3 and Table 4.2, where lifespan graphs and Kaplan-Meier statistics are shown respectively. While knockdown of most genes resulted in a minor increase in survival of the population, the genes with the largest increase in lifespan are all related to translation. Just as in the lifespan assay, it is seen that

A

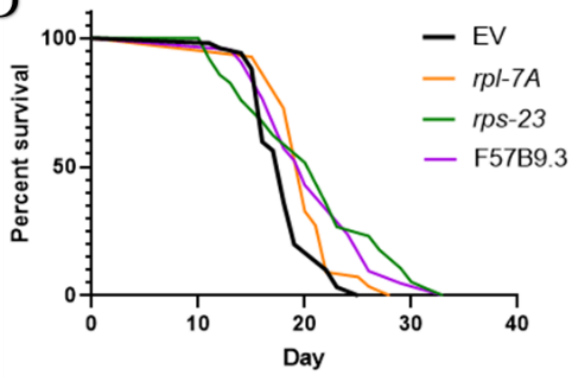
Gene Name	Lifespan (Days)
EV	18.01
ifg-1	↑ 24.30
rpl-7A	↑ 20.24
rps-6	15.27
BMS1	↑ 21.27
rps-23	↑ 20.60
F57B9.3	↑ 20.73
Y39E4B.1	15.30
H12I13.2	18.49
F47B3.6	↓ 12.78
nhr-61	19.50
F46F11.1	↓ 10.91
csp-2	18.74

C



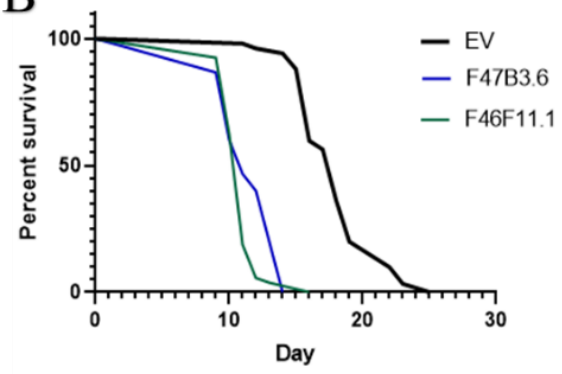
Minimal change in lifespan

D



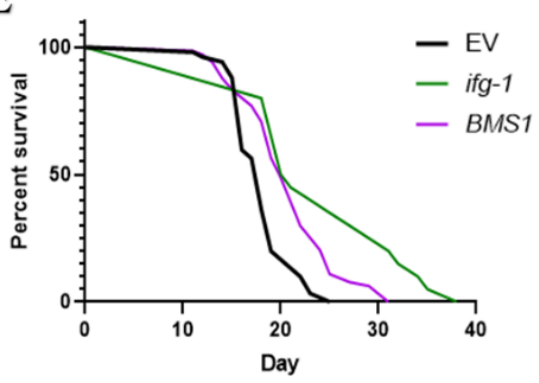
↑ Increase in lifespan

B



↓ Strong decrease in lifespan

E



↑ Strong increase in lifespan

Figure 4.2: Lifespan curves for the RNAi characterization lifespan assay

Lifespan curves for the first trial of the 12 genes of interest, sorted by effect on lifespan compared to the control (EV) curve. **A)** The twelve genes chosen for this assay. The first half (bold) are well-characterized and researched prior, the second half (not bold) are novel genes with little to no prior research. Results of the lifespan can be seen schematically beside each gene name, where green arrows represent an increase in lifespan, red arrows represent a decrease in lifespan, and no arrow represents statistically insignificant change. **B)** Decreased lifespan with respect to the control population; **C)** Minimal change with respect to the control population; **D)** Slightly extended lifespan with respect to the control population; **E)** Highly extended lifespan with respect to the control population. Approximately 40 animals were used per condition per trial. Changes with respect to lifespan were categorized qualitatively. EV: wild-type; *rps-6*: ribosomal protein, small subunit 6; *csp-2*: caSPase-2; *nhr-61*: nuclear hormone receptor 61; *rpl-7A*: ribosomal protein, large subunit 7A; *rps-23*: ribosomal protein, small subunit 23; *ifg-1*: initiation factor 4G (eIF4G); *BMS1:* ribosome biogenesis protein 1.

Table 4.1: Gene function lifespan assay statistical data

Statistical analysis of the four trials completed for the Lifespan Assay. Data was analyzed by the Online Application for Survival Analysis 2 (OASIS 2) [62] using Kaplan-Meier estimator and log-rank tests. EV: wild-type; *rps-6*: ribosomal protein, small subunit 6; *csp-2*: caSPase-2; *nhr-61*: nuclear hormone receptor 61; *rpl-7A*: ribosomal protein, large subunit 7A; *rps-23*: ribosomal protein, small subunit 23; *ifg-1*: initiation factor 4G (eIF4G); *BMS1:* ribosome biogenesis protein 1.

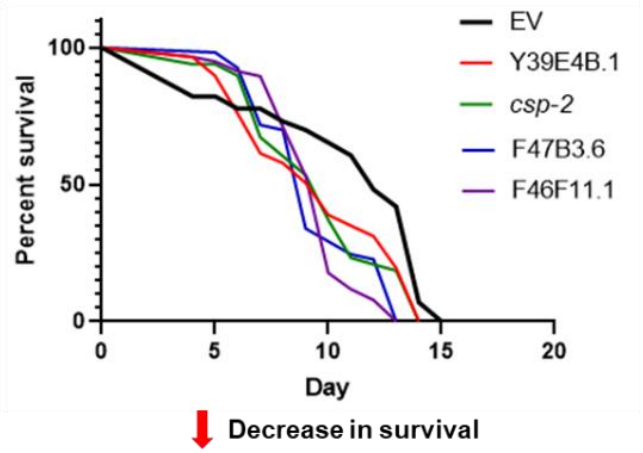
Gene Name	Animals tested	Animals dead	Animals censored	Mean lifespan (Days)	% Mean Lifespan	p value
Lifespan Trial 1						
EV	52	32	20	18.01 ± 0.50	~	~
ifg-1	20	20	0	24.30 ± 1.47	134.93	0.0001
rpl-7A	55	55	0	20.24 ± 0.36	112.38	0.0006
rps-6	47	45	2	15.27 ± 0.57	84.79	0.0039
BMS1	46	43	3	21.27 ± 0.70	118.10	0.0007
rps-23	70	58	12	20.60 ± 0.82	114.38	0.0169
F57B9.3	22	21	1	20.73 ± 1.10	115.10	0.015
Y39E4B.1	63	61	2	15.30 ± 0.24	84.95	5.2E-07
H121I3.2	75	69	6	18.49 ± 0.46	102.67	0.4239
F47B3.6	9	7	2	12.78 ± 0.38	70.96	0
nhr-61	33	32	1	19.50 ± 0.83	108.27	0.0426
F46F11.1	53	53	0	10.91 ± 0.18	60.58	0
csp-2	22	21	1	18.74 ± 0.53	104.05	0.3654
Lifespan Trial 2						
EV	21	20	1	18.79 ± 0.67		~
ifg-1	47	47	0	21.38 ± 0.68	113.78	0.0069
rpl-7A	87	84	3	19.31 ± 0.31	102.77	0.3884
rps-6	65	56	9	17.91 ± 0.37	95.32	0.3034
BMS1	30	30	0	20.87 ± 1.25	111.07	0.2852
rps-23	66	61	5	19.61 ± 0.60	104.36	0.2449
F57B9.3	26	26	0	20.77 ± 1.10	110.54	0.0858
Y39E4B.1	75	69	6	15.44 ± 0.25	82.17	2.7E-06
H121I3.2	87	81	6	17.54 ± 0.26	93.35	0.0135
F47B3.6	41	36	5	12.46 ± 0.17	66.31	0
nhr-61	23	22	1	18.80 ± 0.38	100.05	0.4144
F46F11.1	50	50	0	11.32 ± 0.09	60.24	0
csp-2	37	37	0	16.27 ± 0.45	86.59	0.0038
Lifespan Trial 3						
EV	19	13	6	15.06 ± 1.15		~
ifg-1	65	63	2	21.76 ± 0.83	144.49	0.000012
rpl-7A	47	46	1	18.03 ± 0.50	119.72	0.0162
rps-6	31	28	3	16.80 ± 0.54	111.55	0.5849
BMS1	15	14	1	17.36 ± 1.66	115.27	0.2768

rps-23	65	62	3	19.90 ± 0.61	132.14	0.0002
F57B9.3	43	38	5	16.82 ± 0.67	111.69	0.2188
Y39E4B.1	50	40	10	14.29 ± 0.48	94.89	0.4704
H121I3.2	54	48	6	16.41 ± 0.36	108.96	0.5881
F47B3.6	58	47	11	12.60 ± 0.21	83.67	0.1372
nhr-61	44	42	2	19.43 ± 0.64	129.02	0.0012
F46F11.1	36	35	1	9.83 ± 0.16	65.27	1E-07
csp-2	22	21	1	14.04 ± 0.50	93.23	0.259
Lifespan Trial 4						
EV	32	26	6	19.87 ± 0.56		~
ifg-1	40	35	5	19.68 ± 0.53	99.04	0.8793
rps-6	67	54	13	19.23 ± 0.40	96.78	0.5266
BMS1	36	33	3	19.92 ± 0.57	100.25	0.429
F57B9.3	44	41	3	19.70 ± 0.74	99.14	0.1125
Y39E4B.1	22	20	2	17.25 ± 0.79	86.81	0.0091
F47B3.6	70	63	7	12.45 ± 0.16	62.66	0
nhr-61	30	27	3	19.26 ± 0.53	96.93	0.4192
F46F11.1	39	38	1	10.21 ± 0.21	51.38	0

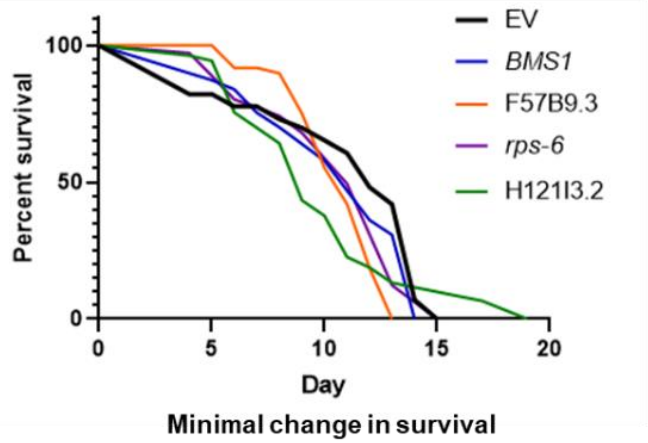
A

Gene Name	Lifespan (Days)
EV	10.86
ifg-1	↑ 11.75
rpl-7A	↑ 12.38
rps-6	10.50
BMS1	10.59
rps-23	↑ 12.58
F57B9.3	10.56
Y39E4B.1	↓ 9.56
H12I13.2	9.89
F47B3.6	↓ 9.47
nhr-61	↑ 12.31
F46F11.1	↓ 9.34
csp-2	↓ 9.55

B



C



D

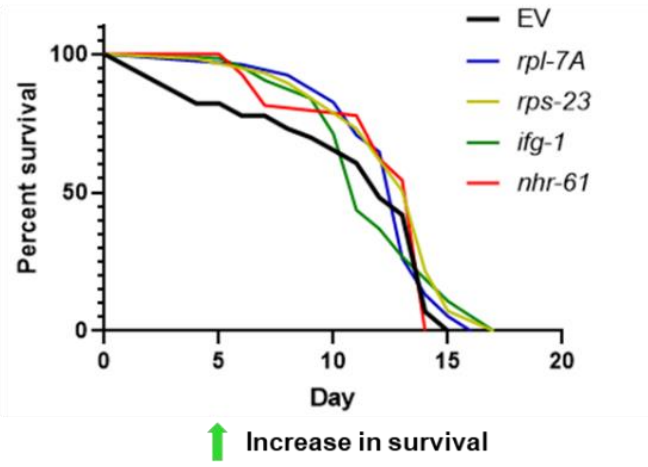


Figure 4.3: Cadmium Sensitivity Assay

Cadmium survival curves for the third trial of the 12 genes of interest, sorted by effect on lifespan compared to the control (EV) curve. **A:** The twelve genes chosen for this assay. The first half (bold) are well-characterized and researched prior, the second half (not bold) are novel genes with little to no prior research. Results of the lifespan can be seen schematically beside each gene name, where green arrows represent an increase in lifespan, red arrows represent a decrease in lifespan, and no arrow represents statistically insignificant change. **B:** Decreased lifespan with respect to the control population; **C:** Minimal change with respect to the control population; **D:** Extended lifespan with respect to the control population. Approximately 40 animals were used per condition per trial. Changes with respect to lifespan were categorized qualitatively. EV: wild-type; *rps-6*: ribosomal protein, small subunit 6; *csp-2*: caSPase-2; *nhr-61*: nuclear hormone receptor 61; *rpl-7A*: ribosomal protein, large subunit 7A; *rps-23*: ribosomal protein, small subunit 23; *ifg-1*: initiation factor 4G (eIF4G); *BMS1:* ribosome biogenesis protein 1.

Table 4.2: Gene function cadmium sensitivity assay statistical data

Statistical analysis of the four trials completed for the cadmium sensitivity assay. Data was analyzed by the Online Application for Survival Analysis 2 (OASIS 2) [62] using the Kaplan-Meier estimator and long-rank tests. EV: wild-type; *rps-6*: ribosomal protein, small subunit 6; *csp-2*: caSPase-2; *nhr-61*: nuclear hormone receptor 61; *rpl-7A*: ribosomal protein, large subunit 7A; *rps-23*: ribosomal protein, small subunit 23; *ifg-1*: initiation factor 4G (eIF4G); *BMS1:* ribosome biogenesis protein 1.

Gene Name	Animals tested	Animals dead	Animals censored	Mean survival (Days)	% Mean survival	p value
Survival Trial 1						
EV	87	41	46	10.20 ± 0.48	~	~
ifg-1	57	55	2	12.20 ± 0.30	119.61	0.0394
rpl-7A	40	36	4	13.87 ± 0.44	135.98	0.000033
rps-6	45	42	3	14.23± 0.37	139.51	0.000001
BMS1	48	44	4	10.25 ± 0.37	100.49	0.0883
rps-23	43	42	1	16.26 ± 0.49	159.41	0
F57B9.3	63	53	10	12.22 ± 0.30	119.80	0.0754
Y39E4B.1	68	59	9	12.53 ± 0.28	122.84	0.0044
H121I3.2	53	52	1	9.12 ± 0.34	89.41	0.0216
F47B3.6	59	53	6	10.94 ± 0.24	107.25	0.8535
nhr-61	54	46	8	11.28 ± 0.47	110.59	0.1997
F46F11.1	61	54	7	10.89 ± 0.20	106.76	0.4364
csp-2	59	56	3	10.85 ± 0.31	106.37	0.9719
Survival Trial 2						
EV	68	39	29	9.82 ± 0.44	~	~
ifg-1	39	36	3	11.30 ± 0.47	115.07	0.0562
rpl-7A	41	41	0	11.98 ± 0.47	122.00	0.0006
rps-6	42	37	5	11.25 ± 0.42	114.56	0.1831
BMS1	46	37	9	10.16 ± 0.47	103.46	0.8452
rps-23	41	36	5	13.00 ± 0.51	132.38	0.0005
F57B9.3	41	35	6	10.98 ± 0.42	111.81	0.2728
Y39E4B.1	47	42	5	11.99 ± 0.37	122.10	0.0137
H121I3.2	55	49	6	10.10 ± 0.32	102.85	0.6382
F47B3.6	59	53	6	10.25 ± 0.15	104.38	0.4775
nhr-61	19	16	3	8.44 ± 0.59	85.95	0.0623
F46F11.1	40	31	9	10.67 ± 0.31	108.66	0.9914
csp-2	31	27	4	10.79 ± 0.50	109.88	0.7092
Survival Trial 3						
EV	73	45	28	10.86 ± 0.47	~	~
ifg-1	65	50	15	11.75 ± 0.36	108.20	0.6947
rpl-7A	51	45	6	12.29 ± 0.32	113.17	0.9334
rps-6	36	32	4	10.50 ± 0.55	96.69	0.0978
BMS1	66	42	24	10.75 ± 0.41	98.99	0.4776

rps-23	59	34	25	12.58 ± 0.42	115.84	0.0913
F57B9.3	51	44	7	10.64 ± 0.28	97.97	0.0011
Y39E4B.1	29	24	5	9.56 ± 0.62	88.03	0.0316
H121I3.2	54	48	6	9.89 ± 0.54	91.07	0.004
F47B3.6	55	53	2	9.47 ± 0.32	87.20	0.000024
nhr-61	30	14	16	12.13 ± 0.54	111.69	0.1886
F46F11.1	59	51	8	9.34 ± 0.26	86.00	2.5E-06
csp-2	49	37	12	9.65 ± 0.43	88.86	0.0087
Survival Trial 4						
EV	60	25	35	10.31 ± 0.59	~	~
Y39E4B.1	26	17	9	11.18 ± 0.54	108.44	0.6625
nhr-61	56	28	28	10.06 ± 0.66	97.58	0.9617

suppression of translation via knockdown of a single gene related to protein synthesis causes an increase in survival of a population of *C. elegans* under cadmium stress, while knockdown of genes unrelated to translation had little effect on the cadmium sensitivity assay. Overall, results here show that suppression of translation is sufficient for lifespan extension as well as for protection against environmental stresses that affect RNA splicing, such as cadmium.

4.3: Obtaining a *C. elegans* strain with mutation to *ifg-1*

I then sought to obtain a stable strain of *C. elegans* carrying an *ifg-1* mutation, thus eliminating the need for RNAi knockdown *E. coli* feeding before each experiment, which is known to produce variabilities across animals and between experiments. Working with a stable mutant strain is also advantageous as RNAi can be used in conjunction, which permits studying potential epistatic relationship between two genes. The *ifg-1* mutant was chosen as it was the only gene with a viable mutant with defect in translation from list of 64 genes obtained from the genome-wide RNAi screen. To start, I obtained the KX54 strain from the CGC that was initially created by Dr. Brett Keiper's lab at East Carolina University (Table 3.1, Materials and Methods). The *ifg-1* transcript reading frame in the KX54 strain is disrupted by a *Mos* transposon insertion that results in premature termination of the mRNA transcript, as depicted in Figure 4.4A. This strain came with the desired *ifg-1* mutation, however it also came with a secondary transgenes expressing CED-1 protein tagged to GFP that is unrelated to my study. Before research could be done with this strain, this unrelated GFP tag must be removed. To achieve this, I performed genetic crosses between KX54 hermaphrodites and wild-type N2 males to remove the CED-1::GFP transgenes while still retaining the *ifg-1* mutation. In this cross, a single F1 hermaphrodite offspring are selected that are heterozygous for the *ifg-1* mutation (*ifg-1*/wildtype) and the CED-1::GFP transgene (GFP/no-GFP). Given that the *ifg-1* mutation and CED-1::GFP are not linked they will segregate independently, and the odds of recovering a (*ifg-1/ifg-1* ; no-GFP/no-GFP) is 1 in 16 (1 in 4 chance of each mutation). 30 F2 offspring were singled onto individual plates and allowed to lay F3 progenies, after which the F2 hermaphrodite is removed from the plate and used to perform single worm genotype PCR. A triple primer strategy is used to genotype for the homozygous *ifg-1* mutation (Figure 4.4A) with an expected band size of 453 bp for *ifg-1* mutation and 353 bp for wildtype. Out of these 30 F2 offspring, only three were

found to have the desired homozygous *ifg-1* mutation as indicated by the larger band product of 454 bp shown in Figure 4.4B. Plate 21 was chosen out of these three based on the health of the population, and this population was then deemed the *ifg-1* mutant strain and used for all further experiments that include the *ifg-1* mutant strain in this research.

4.4: *ifg-1* lifespan regulation closely linked to RNA splicing

In order to determine the effect of *ifg-1* knockdown and cadmium exposure on a transcriptome level, RNA extraction and next generation RNA sequencing was performed on N2 and *ifg-1* mutants with and without cadmium exposure. This allowed for an understanding of the global effect of *ifg-1* knockdown on gene expression and RNA splicing/mis-splicing with and without cadmium exposure. To start, L4 stage N2 or *ifg-1* mutants were moved to plates with or without 300 μM cadmium chloride, as illustrated in Figure 4.5A. After a 24 hour exposure period, RNA was extracted via sonication and by following the Invitrogen™ PureLink™ RNA Mini Kit manual. This resulted in a small amount of concentrated RNA extracted from a large population of worms. After normalizing all samples to 120 ng/μl, samples were ran on a 2% agarose gel as a method of checking RNA integrity; the results of this gel can be seen in Figure 4.5B, where it is seen that all 16 samples (four replicates per condition) produced the intact 18S and 28S ribosomal RNA. If samples had been degraded, the gel would have produced a "smear" of RNA in each well instead of distinct bands. Before an aliquot of the 16 RNA samples were sent to Novogene for RNA-sequencing, two PCR assays were performed first to ensure the samples were accurately prepared. Prior to PCR, samples were treated with DNAse to remove any DNA in each sample, then converted to cDNA for further use. First, gel-based PCR was performed to measure RNA splicing of the endogenous ret-1 gene to confirm the changes in GFP pattern from the KH2235 worms used in the genome-wide RNAi screen. The *ret-1* primer binds to exon 4 and 6 and produces two PCR products with the small product indicating the exon skipped isoform. PCR was performed using primers for the *ret-1* gene to confirm the proper amount of skipped vs unskipped *ret-1* was present in each sample; these data can be seen in Figure 4.5C and quantified in Figure 4.5D. Exposure to cadmium significantly increased the

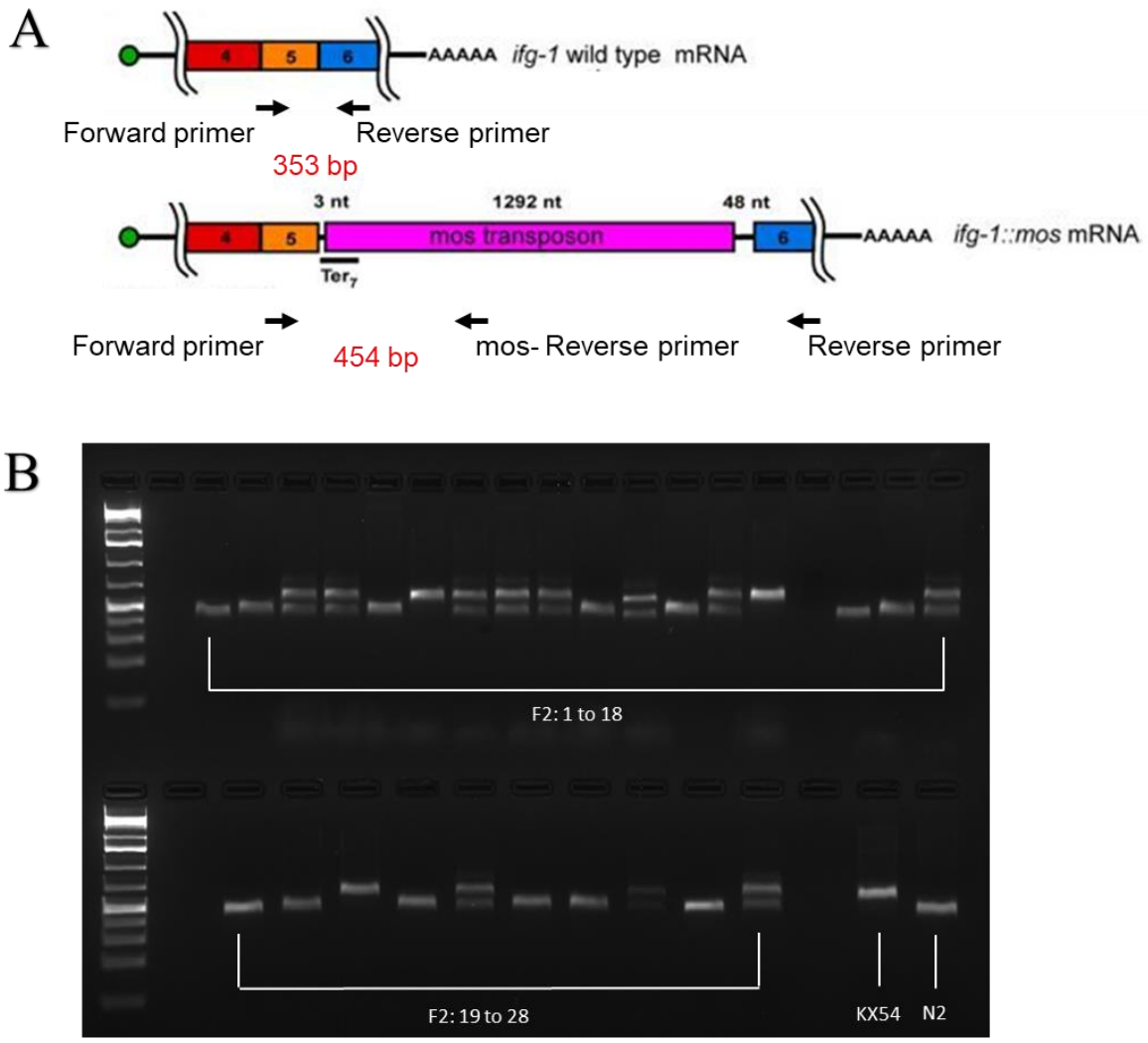
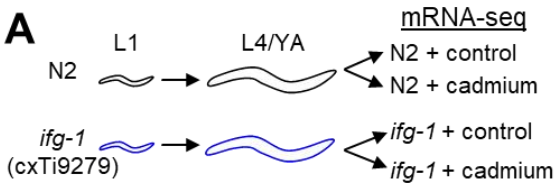
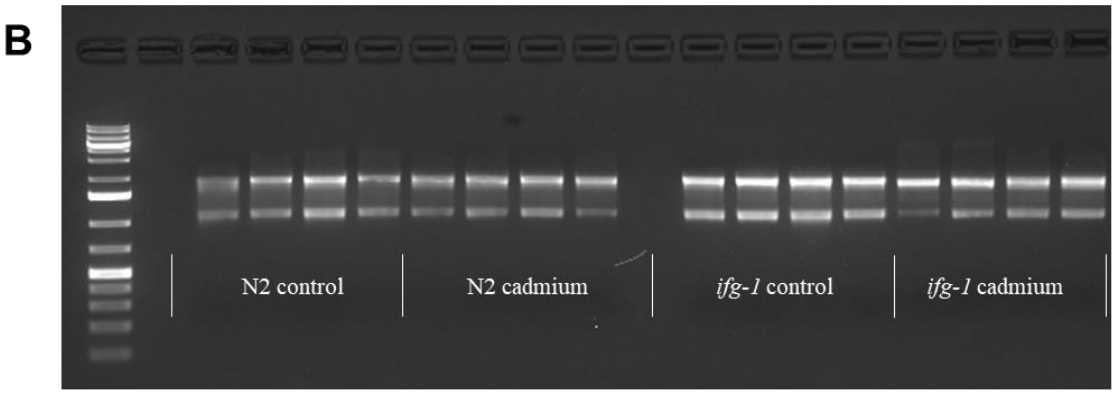
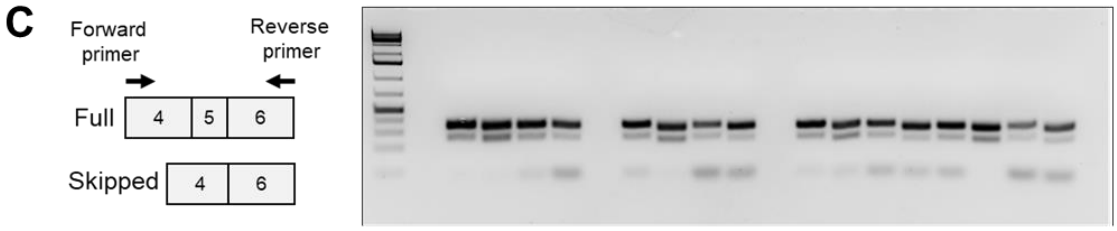


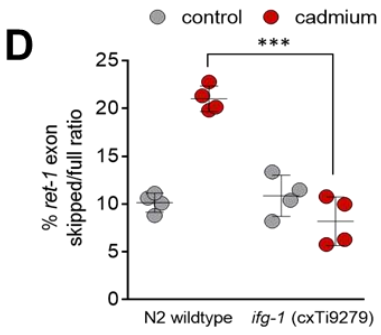
Figure 4.4: KX54 x N2 genetic cross

A) Diagram of the *ifg-1* mutant strain used in this study, where the upper sequence represents the wild-type *ifg-1* gene and the lower sequence represents the loss-of-function *ifg-1* strain. The insertion of a *Mos* transposon disrupt the reading frame of *ifg-1*. Expected PCR amplicon sizes are represented in red, and primers and direction are represented with black arrows. Figure adapted from [59]. **B)** Gel electrophoresis of the 30 F2 offspring after single worm lysis and PCR, where the first 28 lane correspond to the 28 plates used in this experiment, followed by one lane of KX54 worms and one lane of N2 worms as controls. A 2% gel was separated for 45 minutes at 100V on a PowerPac Basic electrophoresis chamber (Bio-Rad). Gels were imaged using a ChemiDoc MP Imaging System (Bio-Rad).


A
mRNA-seq
L1
L4/YA
N2
N2 + control
N2 + cadmium
ifg-1
(cxTi9279)
ifg-1 + control
ifg-1 + cadmium


B
N2 control
N2 cadmium
ifg-1 control
ifg-1 cadmium


C
Forward primer
Reverse primer
Full
4
5
6
Skipped
4
6


D
control
cadmium
***
% ret-1 exon skipped/full ratio
25
20
15
10
5
0
N2 wildtype
ifg-1 (cxTi9279)

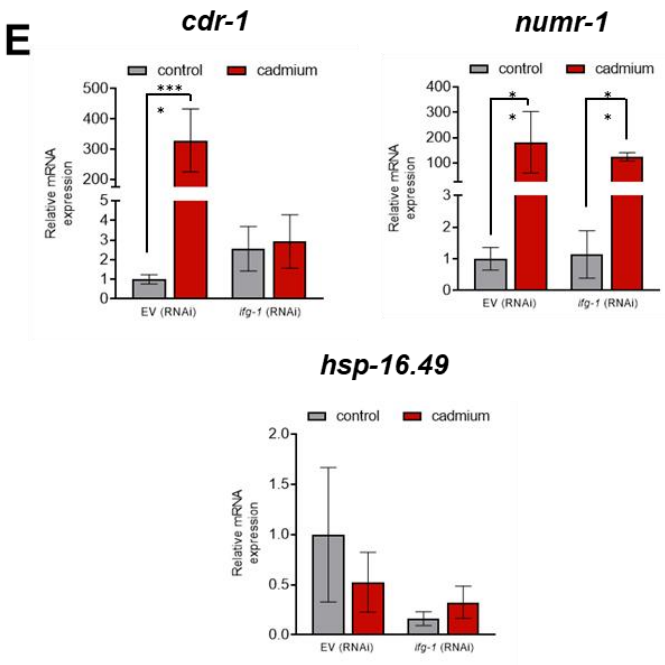

E
cdr-1
numr-1
hsp-16.49
control
cadmium
Relative mRNA expression
EV (RNAi)
ifg-1 (RNAi)

Figure 4.5: RNA Extraction of *ifg-1* mutants on cadmium

Results from the RNA extraction of *ifg-1* and N2 worms exposed to cadmium, including extraction and quality control. **A)** Simplified workflow of the RNA extraction; N2 (black) and *ifg-1* mutant (blue) worms are grown to the L4 stage, then half of each population is exposed to cadmium. After an additional 24 hours, RNA was extracted from the animals. **B)** Agarose gel results of the quality control check post-RNA extraction. Lane 1: 1 kb ladder; lanes 2-5: N2 control; lane 6-8: N2 cadmium, lane 9-12: *ifg-1* control; lane 13-16: *ifg-1* cadmium. The expected banding pattern included two distinct bands at 5,000 and 2,000, and an unsuccessful RNA extraction would be indicated by a "smear" of RNA in the lane instead. A 2% gel was ran for 30 minutes at 120V on a PowerPac Basic electrophoresis chamber (Bio-Rad). Gels were imaged using a ChemiDoc MP Imaging System (Bio-Rad). **C)** Figure depicting the sequence of each band in an agarose gel of KH2235 worms, with a corresponding agarose gel beside. Each four lanes is a repeating sequence of: N2 control, N2 cadmium, *ifg-1* control, and *ifg-1* cadmium. Four replicates of each condition is shown. A 2% gel was ran for 30 minutes at 120V on a PowerPac Basic electrophoresis chamber (Bio-Rad). Gels were imaged using a ChemiDoc MP Imaging System (Bio-Rad). **D)** Graph of relative band intensity of each of the samples analyzed in the gel in Figure 4.5C, where the X axis represents each sample and the Y axis represents the percent of skipped/full ratio of the *ret-1* reporter gene. One-way ANOVA was used for significance testing. *** indicates $P < 0.001$. **E)** qPCR results of the N2 and *ifg-1* populations with and without cadmium exposure against three cadmium response genes: *cdr-1, numr-1*, and *hsp-16.49.* **** indicates $P < 0.0001$, ** indicates $P < 0.01$. Four replicates were used, with each replicate having a sample size of 200 to 300 worms. Statistical analysis is available in Appendix C.

proportion of the skipped *ret-1* isoform, and this increase in exon skipping is not observed in the *ifg-1* mutant, consistent with the retention of GFP signals observed in the RNAi screen.

Next, qPCR was performed on the converted cDNA to confirm the response to cadmium under each condition through the use of three known cadmium response genes: cadmium responsive 1 (*cdr-1);* nuclear localized metal responsive 1 *(numr-1);* and heat shock protein 16.49 (*hsp-16.49).* Alongside these response genes we tested the "housekeeping gene" ribosomal protein large subunit 2 (*rpl-2*), whose expression does not change due to cadmium exposure; this gene acts as a control for the qPCR and allows us to see background variance between samples that is not due to cadmium exposure. As seen graphically in Figure 4.5E and statistically via two-way ANOVA in Appendix C, a significant change between the expression of *cdr-1* in N2 worms exposed to cadmium was seen, but the expression of this gene was drastically reduced in *ifg-1* knockdown worms exposed to cadmium due to *ifg-1* protecting against cadmium based stresses. For *numr-1* expression, there was a significant difference between control worms and worms exposed to cadmium, however no difference between N2 and *ifg-1* knockdown worms was observed. And for *hsp-16.49*, no significant differences were observed between any conditions.

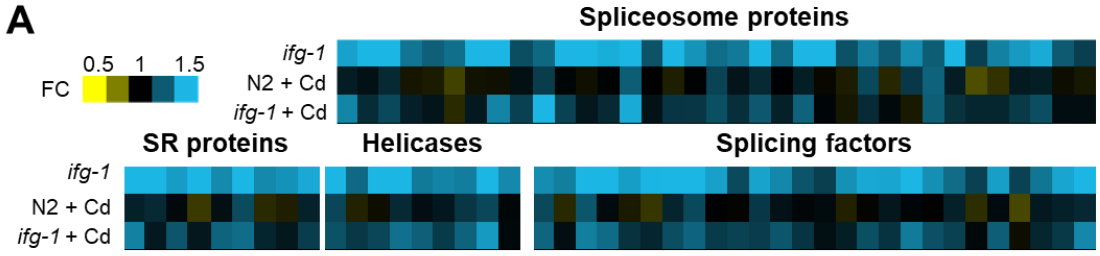
Samples were sent to Novogene where a poly-A tail enriched library were prepared and sequenced followed by bioinformatics analysis (Appendix A). Once data was obtained, I was able to fully appreciate the genome-wide effects that *ifg-1* knockdown had on cadmium exposed worms. Firstly, it was seen that *ifg-1* loss of function worms upregulates over 80 genes known to regulate RNA splicing, and is able to moderately retain upregulation after cadmium exposure The heatmap featured in Figure 4.6A, which compares the expression of genes against wild-type N2 worms without cadmium, shows that an assortment of RNA splicing genes are moderately upregulated in *ifg-1* mutants as well as in *ifg-1* mutants exposed to cadmium, albeit less so than in the non-cadmium *ifg-1* population. Next, the data also elucidated that *ifg-1* loss of function reduces the amount of cadmium-induced alternative splicing events by ~50%. These data are presented in Figure 4.6B, which breaks down the total significant AS events per strain (N2 in black, *ifg-1* mutants in blue) into the five main categories: exon skipping, mutually exclusive exons, alternative 5' or 3' splice site, or intron retention. While specific variances between N2 and *ifg-1* are seen in each category (AS type, upregulation or downregulation), the total amount of significant events decreases by half when comparing N2 to *ifg-1* mutants. The relationship

between alternative splicing and the conditions studied can also be seen in Figure 4.6C, which plots the alternative splicing of each isoform of a gene for N2 worms exposed to cadmium against *ifg-1* worms exposed to cadmium. It can be seen that there is a negative and significant correlative relationship (-0.37) between the alternative splicing of an isoform between the wildtype N2 strain and the *ifg-1* mutant after cadmium exposure. When assessing individual alternative splicing events, it revealed that a negative correlation existed for exon skipping, intron retention, and mutually exclusive exon events, but not 5' or 3' alternative splice site events. (Figure 4.6D). A graphical illustration of IncLevel is shown in Figure 4.6E, where IncLevel is equal to the amount of the isoform of interest divided by the total amount of both isoforms; this allows for a normalized comparison between isoforms (Figure 4.6E).

I also sought out to understand the differences in global gene expression changes between *ifg-1* and N2 worms exposed to cadmium through this RNA-seq. I first confirmed that the ifg-1 loss of function strain had decrease in *ifg-1* mRNA levels (57%), suggesting that this is a partial loss of function mutant (Figure 4.7A). I then sought out to characterize the functional categories of genes that had a significant change in expression in the *ifg-1* mutant. Interestingly, genes that were upregulated in the *ifg-1* mutant were highly involved in xenobiotic detoxification (Figure 4.7B), which are also upregulated in wildtype worms exposed to cadmium. Using clustering analysis, I found that N2 worms exposed to cadmium had a striking similarity to *ifg-1* mutants not exposed to cadmium in terms of gene expressional changes. These data can be seen in Figure 4.7C, where blue represents upregulation of a gene with respect to the N2 control population and yellow represents downregulation. In summary, I found that exposure to cadmium with suppression of *ifg-1* results in enhanced RNA splicing via the reduction of AS events in *C. elegans*. As well, I noticed that gene expression changes of wild type worms exposed to cadmium closely mirrored *ifg-1* knockdown worms, indicating that loss of translational capacity may mimic a state of stress in these *ifg-1* mutant worms.

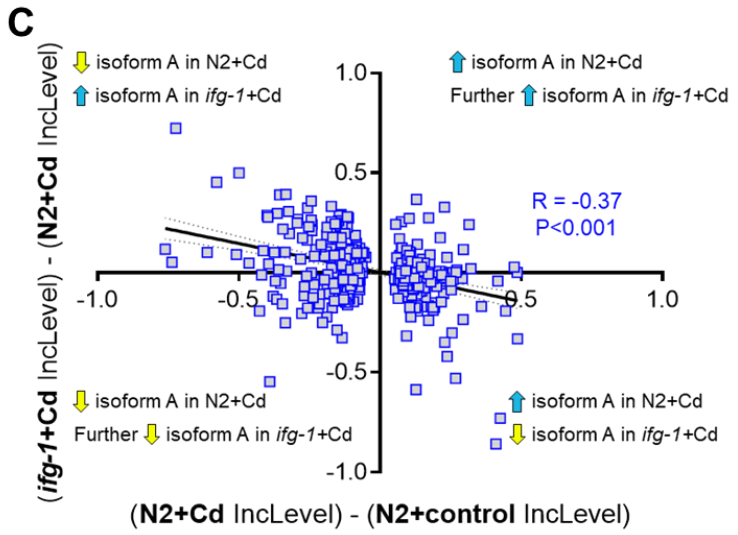
4.5: RNA splicing is required for *ifg-1's* long-lived phenotype

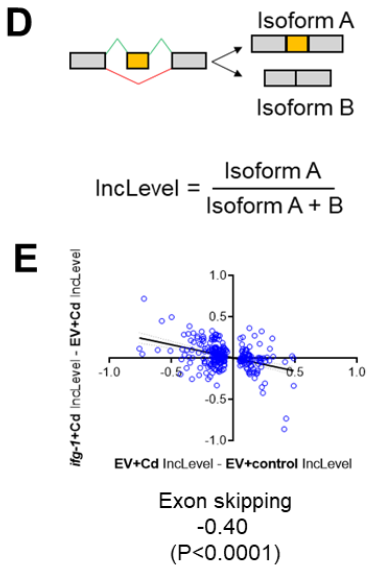
Given that my results indicate that *ifg-1* mutants exhibit a different alternative splicing profile, I next tested whether genes that regulate core RNA splicing functions are required for *ifg-1*'s known long-lived phenotype. This lifespan assay was done with the intent to uncover



B

Alternative splicing event type	Total events detected	Significant events (same events with FDR<0.05 in *ifg-1* + cadmium)	
		IncLevel Difference > 0	IncLevel Difference < 0
Exon skipping	5,940	94 ; (55/94)	183 ; (104/183)
Mutually exclusive exon	633	29 ; (11/29)	35 ; (11/35)
Alternate 5' splice site	286	15 ; (5/15)	20 ; (12/20)
Alternative 3' splice site	497	33 ; (15/33)	10 ; (4/10)
Intron retention	192	12 ; (7/12)	6 ; (4/6)
Total events		183 (93/183)	254 (135/254)





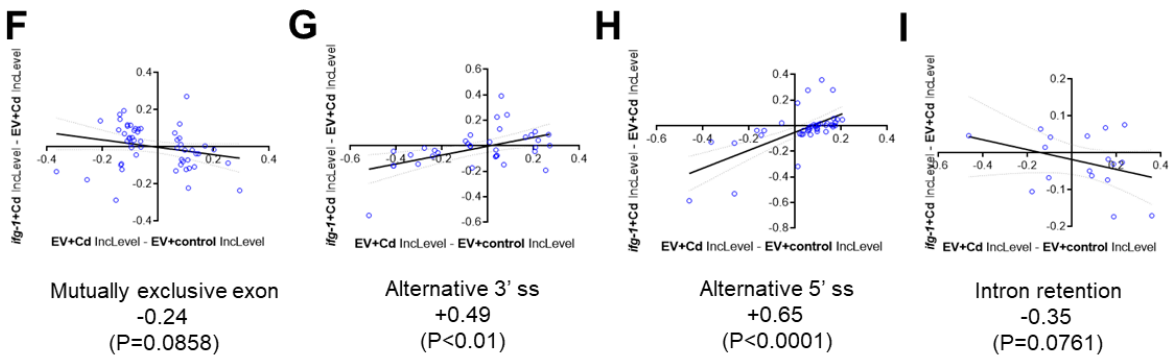


Figure 4.6: Effect of *ifg-1* knockdown on alternative splicing in cadmium exposed *C. elegans* Analysis of data collected from the RNA-seq performed by Novogene. **A)** Differential expression of genes related to RNA splicing in *ifg-1,* N2 cadmium, and *ifg-1* cadmium populations normalized to the N2 control population. Yellow indicates a decrease in expression, while blue indicates an increase in expression. **B)** Alternative splicing events in N2 (black) and *ifg-1* mutant (blue) worms after cadmium exposure. Events are categorized by AS event type and by either increase (IncLevel Difference > 0) or decrease (IncLevel Difference < 0) in expression. FDR < 0.05 indicates events are significantly different. FDR: false discovery rate. **C)** Isoform expression of significantly differentially expressed genes plotted on an XY graph, where X is the IncLevel of N2 control worms subtracted from the IncLevel of N2 cadmium worms, and Y is the IncLevel of N2 cadmium worms subtracted from the IncLevel of *ifg-1* cadmium worms. Each point on the graph represents a significant alternatively spliced gene, and its position on the graph is related to the IncLevel different of that isoform in N2 cadmium or *ifg-1* cadmium exposed populations. **D)** Visual representation of how IncLevel is calculated using the expression of each isoform, with exon skipping as an example AS event.

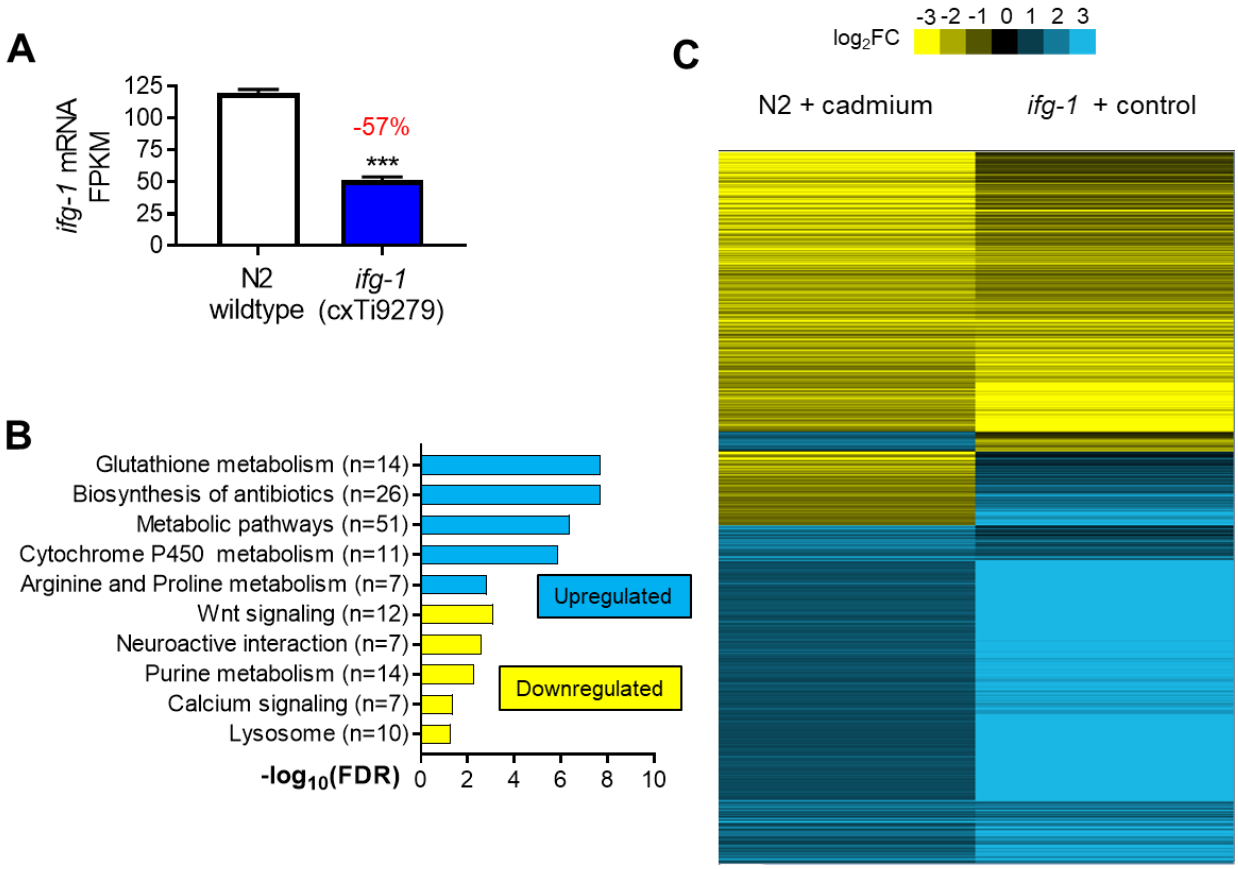


Figure 4.7: Effect of *ifg-1* knockdown on gene expression in cadmium exposed *C. elegans*

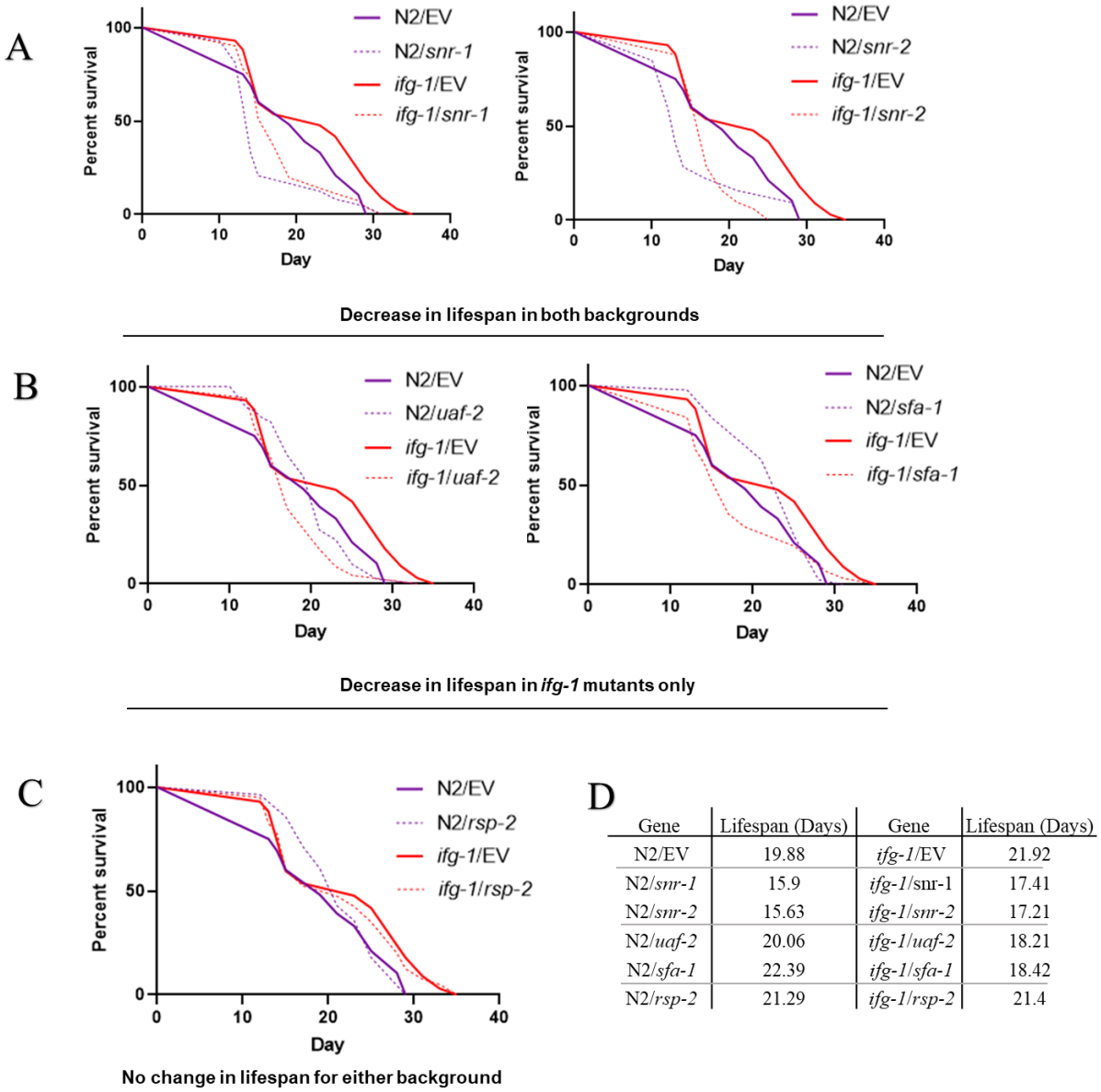
Analysis of data collected from the RNA-seq performed by Novogene. **A)** Bar graph representing the change in expression of *ifg-1* between N2 (white) and *ifg-1* (blue) worms, where Y is the FPKM of *ifg-1* mRNA. A decrease of 57% is seen in *ifg-1* worms when compared to N2 worms. *** indicates a highly significant difference between two samples as determined by Student's t-test P<0.0001. FPKM = Fragments Per Kilobase of transcript per Million mapped reads. **B)** Enrichment analysis of genes significantly upregulated (blue) or downregulated (yellow) in *ifg-1* worms compared to N2 worms exposed to cadmium. FDR: false discovery rate. **C)** Expression level changes when compared to N2 worms not exposed to cadmium for all significant genes between N2 worms exposed to cadmium and *ifg-1* worms not exposed to cadmium. Fold change is compared to the N2 control population. Yellow indicates a decrease in expression, while blue indicates an increase.

which genes are required by *ifg-1* for lifespan regulation by researching *ifg-1*'s interaction with a series of known RNA splicing regulating genes. To do this, *ifg-1* mutants were fed RNAi knockdown *E. coli* targeting a RNA splicing regulator gene of interest, resulting in an *ifg-1* mutant population with a compromised spliceosome. If a gene is required by *ifg-1* for lifespan regulation, I would expect a decrease in *ifg-1* lifespan after the RNAi knockdown.

The seven genes chosen are all well known genes in literature that play a role in RNA splicing. *snr-1* and *snr-2,* small nuclear riboproteins, are both essential for RNA splicing. *uaf-2* enables binding of pre-mRNA at the 3' splice site, while *sfa-1* is involved in AS and branch point splicing. Lastly, *rsp-2, hrpf-1,* and *hrp*-2 are all predicted to enable RNA binding activity. While all seven genes are tested against EV control worms for the lifespan assay, only five (*snr-1, snr-2, sfa-1, uaf-2, rsp-2*) are used in the cadmium sensitivity assay. This decision was due to learning that the population of RNAi *E. coli* targeting for *hrp-2* and *hrpf-1* was determined to actually be targeting a different gene entirely. This was suspected during the lifespan assay and confirmed before the cadmium sensitivity assays began via sequencing the dsRNA plasmids for *hrp-2* and *hrpf-1* bacteria which revealed that they both encoded for a different gene. For this reason, data collected on *hrp-2* and *hrpf-1* during the lifespan assay was not further considered, and is instead presented in Appendix D. The results of this lifespan, as seen graphically in Figure 4.8 and statistically via Kaplan-Meier estimations in Table 4.3, show that four of the five gene knockdowns resulted in a decreased lifespan in *ifg-1* mutants, with only *rsp-2* showing no effect on *ifg-1* lifespan. These lifespan changes are then categorized further by their change in lifespan relative to the N2 and the *ifg-1* mutant fed with EV control RNAi. Knockdown of *snr-1* and *snr-2* shorten the lifespan of both N2 and *ifg-1* mutant, suggesting that the *snr* genes are universally required for longevity. Interestingly, knockdown of *uaf-2* and *sfa-1* reduced only *ifg-1* mutant lifespan, suggesting that these two splicing factors may be directly involved in maintaining the unique alternative splicing pattern of the *ifg-1* mutant. Knockdown of *rsp-2* had no effect on lifespan, suggesting that this splicing factor is dispensable for longevity.

The cadmium sensitivity assay that was performed with the same conditions but saw very little consistent change in survival of the *ifg-1* mutant worms with any of the additional gene knocked down, as seen in Figure 4.9 and Table 4.4. In the N2 background, there were knockdowns that produced an increase in survival of up to 25% to 40% (*sfa-1, uaf-2, rsp-2* in Trial 2), while the

knockdown



Gene	Lifespan (Days)	Gene	Lifespan (Days)
N2/EV	19.88	*ifg-1*/EV	21.92
N2/*snr-1*	15.9	*ifg-1*/snr-1	17.41
N2/*snr-2*	15.63	*ifg-1*/*snr-2*	17.21
N2/*uaf-2*	20.06	*ifg-1*/*uaf-2*	18.21
N2/*sfa-1*	22.39	*ifg-1*/*sfa-1*	18.42
N2/*rsp-2*	21.29	*ifg-1*/*rsp-2*	21.4

Figure 4.8: *ifg-1* mutant background lifespan assay

Lifespan curves of the third trial of *ifg-1* mutants with an additional splicing factor knocked down, categorized by lifespan extension with respect to control. **A)** Genes that, when knocked down, result in a shorter lifespan in both backgrounds (N2 and *ifg-1* mutant); **B)** Genes that, when knocked down, result in a decrease in lifespan in *ifg-1* mutants only; **C)** Genes that, when knocked down, result in either an increase or no change in lifespan in both backgrounds; **D)** Mean lifespans for each condition. Approximately 30 worms were used in each trial. Changes with respect to lifespan were categorized qualitatively. EV: wild-type worms; *snr*: small nuclear ribonucleoprotein; *uaf-2:* U2AF splicing factor; *sfa-1:* splicing factor 1; *rsp-2*: SR protein.

Table 4.3: *ifg-1* mutant background lifespan assay statistical data

Statistical analysis of the five trials completed for the lifespan assay involving *ifg-1* mutants with an additional splicing factor knockdown. Data was analyzed by the Online Application for Survival Analysis 2 (OASIS 2) [62] using the Kaplan-Meier estimator and log-rank tests. EV: wild-type worms; *snr*: small nuclear ribonucleoprotein; *uaf-2:* U2AF splicing factor; *sfa-1:* splicing factor 1; *rsp-2*: SR protein.

Gene Name	Animals tested	Animals dead	Animals censored	Mean lifespan (Days)	% Mean Lifespan	p value
Lifespan Trial 1						
N2/EV	48	34	14	21.34 ± 0.65	~	~
N2/*snr-1*	34	31	3	15.81 ± 0.68	74.09	0.0000016
N2/*snr-2*	50	41	9	18.36 ± 0.74	86.04	0.0093
N2/*sfa-1*	30	26	4	21.81 ± 0.86	102.20	0.2647
N2/*uaf-2*	56	49	7	19.61 ± 0.64	91.89	0.0786
N2/*rsp-2*	23	19	4	21.57 ± 0.82	101.08	0.5776
ifg-1/EV	27	23	4	22.21 ± 1.83	104.08	~
ifg-1/snr-1	39	36	3	17.96 ± 0.49	80.86	0.0486
ifg-1/snr-2	32	30	2	18.97 ± 0.72	85.41	0.1168
ifg-1/sfa-1	43	33	10	20.37 ± 1.38	91.72	0.6155
ifg-1/uaf-2	79	68	11	20.23 ± 0.45	91.09	0.1574
ifg-1/rsp-2	34	28	6	24.03 ± 1.37	108.19	0.6098
Lifespan Trial 2						
N2/EV	36	33	3	19.88 ± 1.02	~	~
N2/*snr-1*	28	25	3	15.9 ± 1.12	79.98	0.0498
N2/*snr-2*	33	32	1	15.63 ± 1.09	78.62	0.0193
N2/*sfa-1*	43	42	1	22.39 ± 0.76	112.63	0.2519
N2/*uaf-2*	59	47	12	20.06 ± 0.68	100.91	0.7227
N2/*rsp-2*	28	26	2	21.29 ± 0.92	107.09	0.7549
ifg-1/EV	44	38	6	21.92 ± 1.18	110.26	~
ifg-1/snr-1	40	37	3	17.41 ± 0.8	79.43	0.1434
ifg-1/snr-2	33	32	1	17.21 ± 0.55	78.51	0.0186
ifg-1/sfa-1	31	31	0	18.42 ± 1.19	84.03	0.5701
ifg-1/uaf-2	35	30	5	18.21 ± 0.88	83.07	0.2264
ifg-1/rsp-2	40	40	0	21.4 ± 1.16	97.63	0.1514
Lifespan Trial 3						
N2/EV	118	104	14	18.77 ± 0.39	~	~
N2/*snr-1*	88	78	10	14.97 ± 0.34	79.75	0
N2/*snr-2*	56	51	5	16.34 ± 0.56	87.05	0.0013
N2/*sfa-1*	116	110	6	19.82 ± 0.40	105.59	0.051
N2/*uaf-2*	26	24	2	17.71 ± 0.79	94.35	0.1746
N2/*rsp-2*	59	52	7	18.95 ± 0.63	100.96	0.6116
ifg-1/EV	125	114	11	19.62 ± 0.61	104.53	~
ifg-1/snr-1	83	78	5	15.67 ± 0.29	79.87	0
ifg-1/snr-2	51	47	4	16.74 ± 0.53	85.32	0.0011
ifg-1/sfa-1	73	66	7	20.75 ± 1.06	105.76	0.1646

ifg-1/uaf-2	95	86	9	19.07 ± 0.66	97.20	0.6788
ifg-1/rsp-2	74	68	6	21.50 ± 0.89	109.58	0.0016
Lifespan Trial 4						
N2/EV	55	47	8	20.24 ± 0.54	~	~
N2/*sfa-1*	74	73	1	20.25 ± 0.80	100.05	0.3703
N2/*uaf-2*	77	72	5	24.65 ± 0.73	121.79	0.0000002
N2/*rsp-2*	53	43	10	23.40 ± 0.74	115.61	0.0004
ifg-1/EV	31	26	5	19.35 ± 1.92	95.60	~
ifg-1/sfa-1	19	19	0	20.95 ± 2.00	108.27	0.3155
ifg-1/uaf-2	25	24	1	20.59 ± 2.05	106.41	0.8345
ifg-1/rsp-2	27	26	1	21.56 ± 1.44	111.42	0.0414

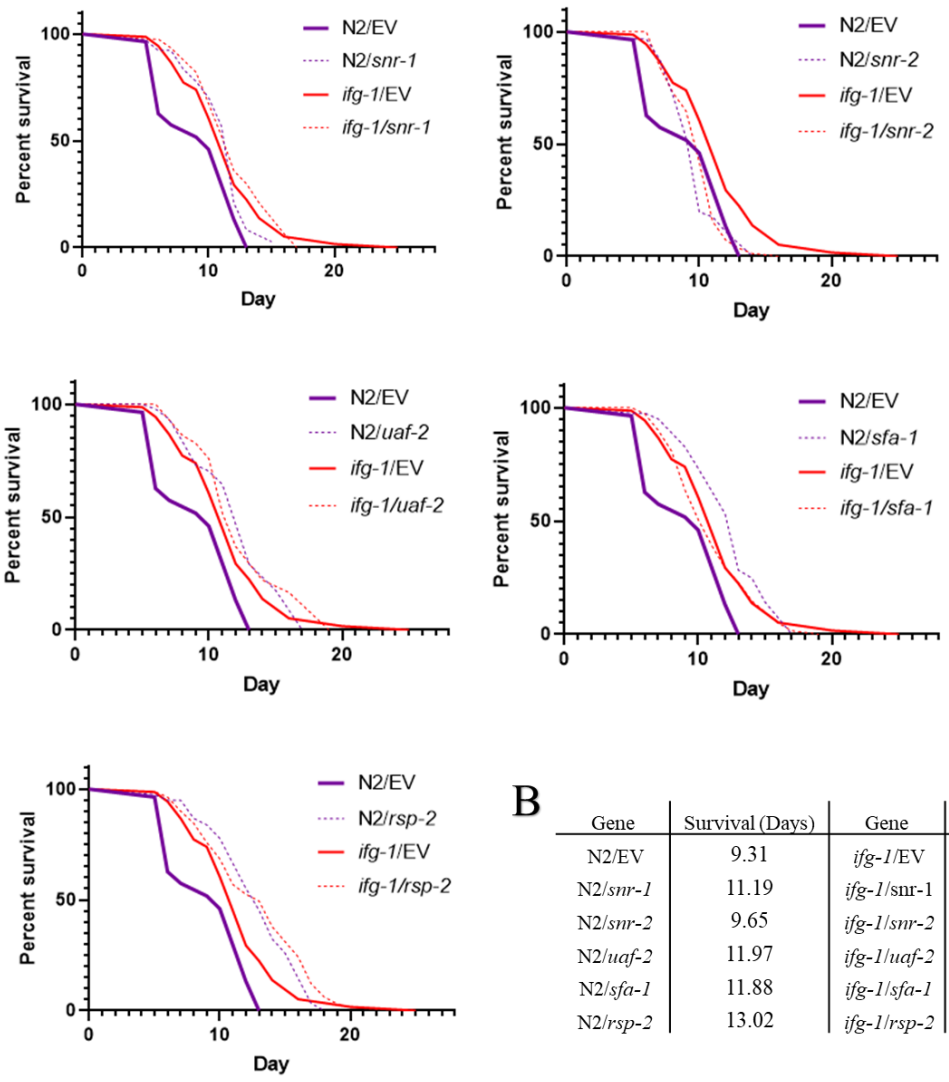
of *snr-2* in Trial 3 saw a decrease in survival of 25%. However, in the *ifg-1* mutant background, the range of survivals was much tighter – a decrease of 16% in Trial 2 when *snr-2* was knocked down, up to an increase of 15% when *rsp-2* was knocked down in Trial 3. The knockdown of each gene resulted in a loss or gain of survival of approximately 10%, compared to the larger range of 20% to 40% for wild-type worms.

Overall, the lifespan data confirms the hypothesis that mutation to *ifg-1* protects RNA splicing fidelity and requires specific components of the RNA splicing machinery to exert its long-lived phenotype. However, RNA splicing does not seem to be required by *ifg-1* for its stress resistance phenotype.

4.6: *ifg-1* signals through the SMA family of proteins

To understand genetic regulators that act downstream of *ifg-1*'s ability to enhance RNA splicing under stress when knocked down, I performed a second RNAi screen with a small set of approximately 1,000 genes of the *C. elegans* genome with a focus on transcription factors and protein kinases and phosphatases. For this RNAi screen, L1 KH2235 worms were initially fed *ifg-1* knockdown RNAi *E. coli* for 48 hours to induce RNA splicing protection, before being exposed to a second RNAi from a sub-library of dsRNA as discussed above. After an additional 48 hours, worms were then exposed to 300 μM cadmium and assessed for their GFP fluorescence. Unlike the prior RNAi screen, 99% of gene knockdowns were able to retain their GFP fluorescence; this was expected and is due to the first knockdown of *ifg-1* resulting in protection of RNA splicing fidelity. While there were no genes that caused complete loss of fluorescence after the second RNAi knockdown, three genes (*sma-2*, *sma-3*, *sma-4*) were identified as genes that, when knocked down in an *ifg-1* depleted background, caused a partial loss of GFP fluorescence after cadmium exposure (Figure 4.10). All three genes identified belong to the same *sma* family of genes, an ortholog of human SMAD proteins. The three genes displayed a varying effect on the fluorescence of the population; *sma-4* had little effect while *sma-3* had the greatest effect.

A



B

Gene	Survival (Days)	Gene	Survival (Days)
N2/EV	9.31	*ifg-1*/EV	11.42
N2/*snr-1*	11.19	*ifg-1*/snr-1	11.92
N2/*snr-2*	9.65	*ifg-1*/*snr-2*	9.61
N2/*uaf-2*	11.97	*ifg-1*/*uaf-2*	11.12
N2/*sfa-1*	11.88	*ifg-1*/*sfa-1*	12.47
N2/*rsp-2*	13.02	*ifg-1*/*rsp-2*	13.02

Figure 4.9: *ifg-1* mutant background cadmium sensitivity assay

Survival curves for the second trial of *ifg-1* mutants with an additional splicing factor knocked down. **A)** Survival curves of each of the seven genes tested against the N2 and *ifg-1* mutant control conditions. Minimal to no change in survival was seen in all knockdown conditions. **B)** Mean lifespans for each condition. Approximately 30 worms were used in each trial. Changes with respect to lifespan were categorized qualitatively. EV: wild-type worms; *snr*: small nuclear ribonucleoprotein; *uaf-2:* U2AF splicing factor; *sfa-1:* splicing factor 1; *rsp-2*: SR Protein.

Table 4.4: *ifg-1* mutant background cadmium sensitivity assay statistical data

Statistical analysis of the three trials completed for the cadmium sensitivity assay involving *ifg-1* mutants with an additional splicing factor knockdown. Data was analyzed by the Online Application for Survival Analysis 2 (OASIS 2) [62] using the Kaplan-Meier estimator and log-rank tests. EV: wild-type worms; *snr*: small nuclear ribonucleoprotein; *uaf-2:* U2AF splicing factor; *sfa-1:* splicing factor 1; *rsp-2*: SR protein.

Gene Name	Animals tested	Animals dead	Animals censored	Mean survival (Days)	% Mean Survival	p value
Survival Trial 1						
N2/EV	58	46	12	10.32 ± 0.38	~	~
N2/*snr-1*	89	73	16	11.18 ± 0.22	108.33	0.3409
N2/*snr-2*	96	79	17	10.04 ± 0.22	97.29	0.2394
N2/*sfa-1*	56	46	10	10.50 ± 0.41	101.74	0.7856
N2/*uaf-2*	70	57	13	11.23 ± 0.29	108.82	0.1777
N2/*rsp-2*	70	59	11	12.49 ± 0.34	121.03	0.0001
ifg-1/EV	38	26	12	9.80 ± 0.52	94.96	0.4895
ifg-1/snr-1	48	37	11	10.76 ± 0.45	109.80	0.2434
ifg-1/snr-2	44	31	13	9.78 ± 0.72	99.80	0.7222
ifg-1/sfa-1	65	46	19	10.72 ± 0.42	109.39	0.1844
ifg-1/uaf-2	35	28	7	10.69 ± 0.51	109.08	0.2585
ifg-1/rsp-2	40	34	6	10.65 ± 0.46	108.67	0.5367
Survival Trial 2						
N2/EV	28	18	10	9.31 ± 0.66	~	~
N2/*snr-1*	69	51	18	11.19 ± 0.30	120.19	0.0342
N2/*snr-2*	62	45	17	9.65 ± 0.29	103.65	0.6648
N2/*sfa-1*	42	30	12	11.97 ± 0.40	128.57	0.0007
N2/*uaf-2*	49	36	13	11.88 ± 0.53	127.60	0.0042
N2/*rsp-2*	41	32	9	13.02 ± 0.56	139.85	0.0001
ifg-1/EV	76	60	16	11.42 ± 0.45	122.66	0.0271
ifg-1/snr-1	83	67	16	11.92 ± 0.35	104.38	0.4303
ifg-1/snr-2	98	85	13	9.61 ± 0.24	84.15	0.0001
ifg-1/sfa-1	76	63	13	11.12 ± 0.38	97.37	0.5742
ifg-1/uaf-2	34	26	8	12.47 ± 0.64	109.19	0.2312
ifg-1/rsp-2	85	69	16	13.02 ± 0.49	114.01	0.0133
Survival Trial 3						
N2/EV	57	32	25	13.17 ± 0.51	~	~
N2/*snr-1*	57	42	15	11.8 ± 0.28	89.60	0.0045
N2/*snr-2*	79	55	24	9.94 ± 0.23	75.47	2.20E-09
N2/*sfa-1*	65	37	28	13.73 ± 0.51	104.25	0.3828
N2/*uaf-2*	57	52	5	12.74 ± 0.38	96.74	0.3798

N2/*rsp-2*	53	38	15	12.25 ± 0.52	93.01	0.3903
ifg-1/EV	58	42	16	12.47 ± 0.57	94.68	0.4163
ifg-1/snr-1	77	52	25	12.94 ± 0.34	103.77	0.6558
ifg-1/snr-2	72	60	12	11.15 ± 0.28	89.41	0.0149
ifg-1/sfa-1	58	42	16	10.72 ± 0.34	85.97	0.0084
ifg-1/uaf-2	67	62	5	10.72 ± 0.42	85.97	0.0188
ifg-1/rsp-2	57	41	16	11.27 ± 0.41	90.38	0.0746

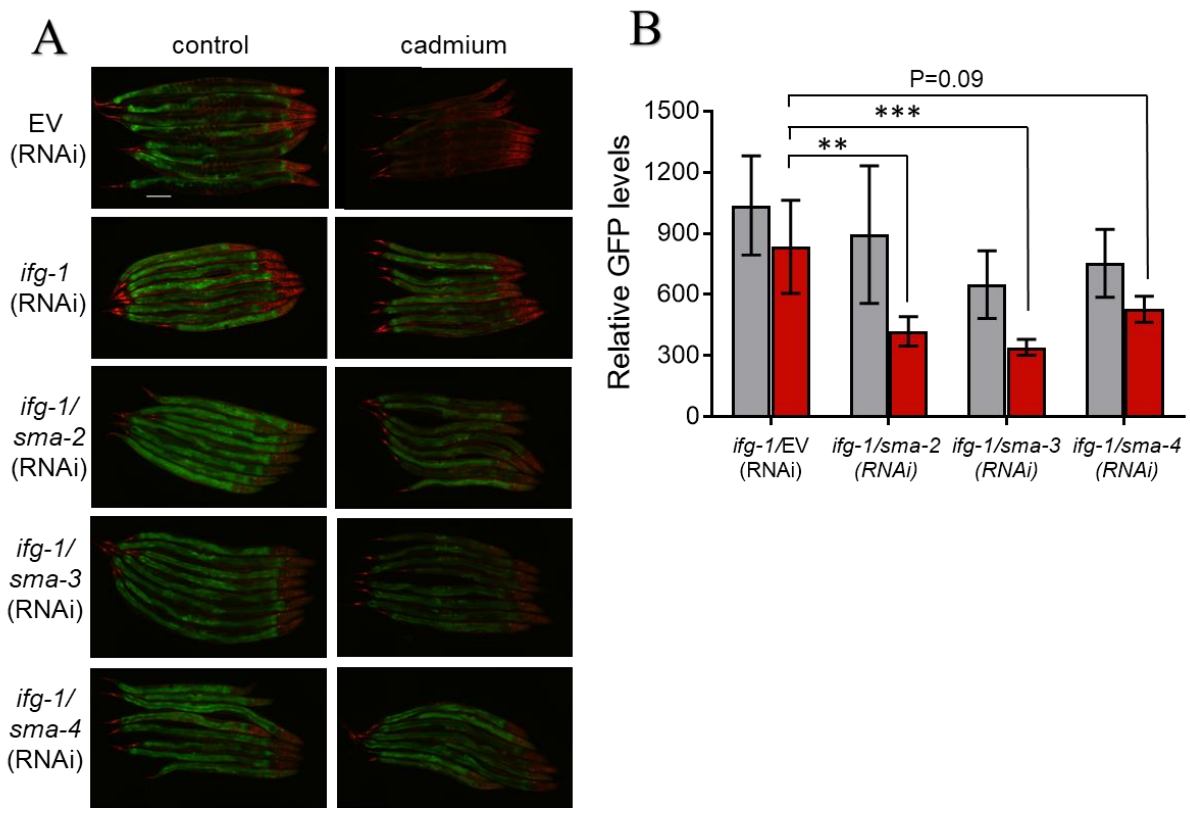


Figure 4.10: Small-scale sub-library RNAi screen

Populations of KH2235 worms with a single or double RNAi knockdown. **A)** Images of KH2235 worms with various genes knocked down via RNAi on control or cadmium conditions. Gene knockdowns are listed to the left of each set of images. Images were taken using a QImaging Retiga R3 (Cairn Research Ltd) attached to a ZEISS Axio Vert.A1 Microscope (Zeiss). **B)** Bar graph of GFP fluorescence of *ifg-1* knockdown worms with or without an additional gene knocked down compared to control worms, where the Y axis is Relative GFP fluorescence and the X axis dictates the knockdown. Grey indicates control conditions and red indicates cadmium conditions. ** indicates $P < 0.001$, *** indicates $P < 0.0001$. Five worms were used per condition.

4.7: SMA-2 regulates majority of *ifg-1* induced genes.

To understand the relationship between *sma* genes and *ifg-1*, I performed meta-analysis of a previously published RNA-seq dataset in a *sma-2* loss of function mutant *sma-2(rax5*) [64]. Surprisingly, there was a large overlap in genes that were upregulated in the *ifg-1* mutant from our experiment but are downregulated in the *sma-2(rax5)* background (Figure 4.11A). The 2,899 genes that overlapped were further analyzed cluster enrichment, which showed a significant enrichment of these 2,899 genes functioning in RNA metabolism including RNA transport and spliceosome (Figure 4.11B). As mentioned earlier, *ifg-1* mutants display upregulation of many genes involved in RNA splicing, and the expression of the majority of the same genes are decreased in the *sma-2(rax5)* background (Figure 4.11C). Generally, genes that are differentially expressed in the *ifg-1* mutant are regulated in the opposite direction in the *sma-2(rax5)* mutant as indicated by the negative correlation value of -0.41 (Figure 4.11D). Overall, the data here suggest that *sma-2* acts as the transcription factor downstream of *ifg-1* to exert its transcriptome changes. This is supported by my results in Figure 4.10 that showed knockdown of *sma-2* abolishes *ifg-1*'s protective effect in resisting cadmium-induced RNA splicing disruptions.

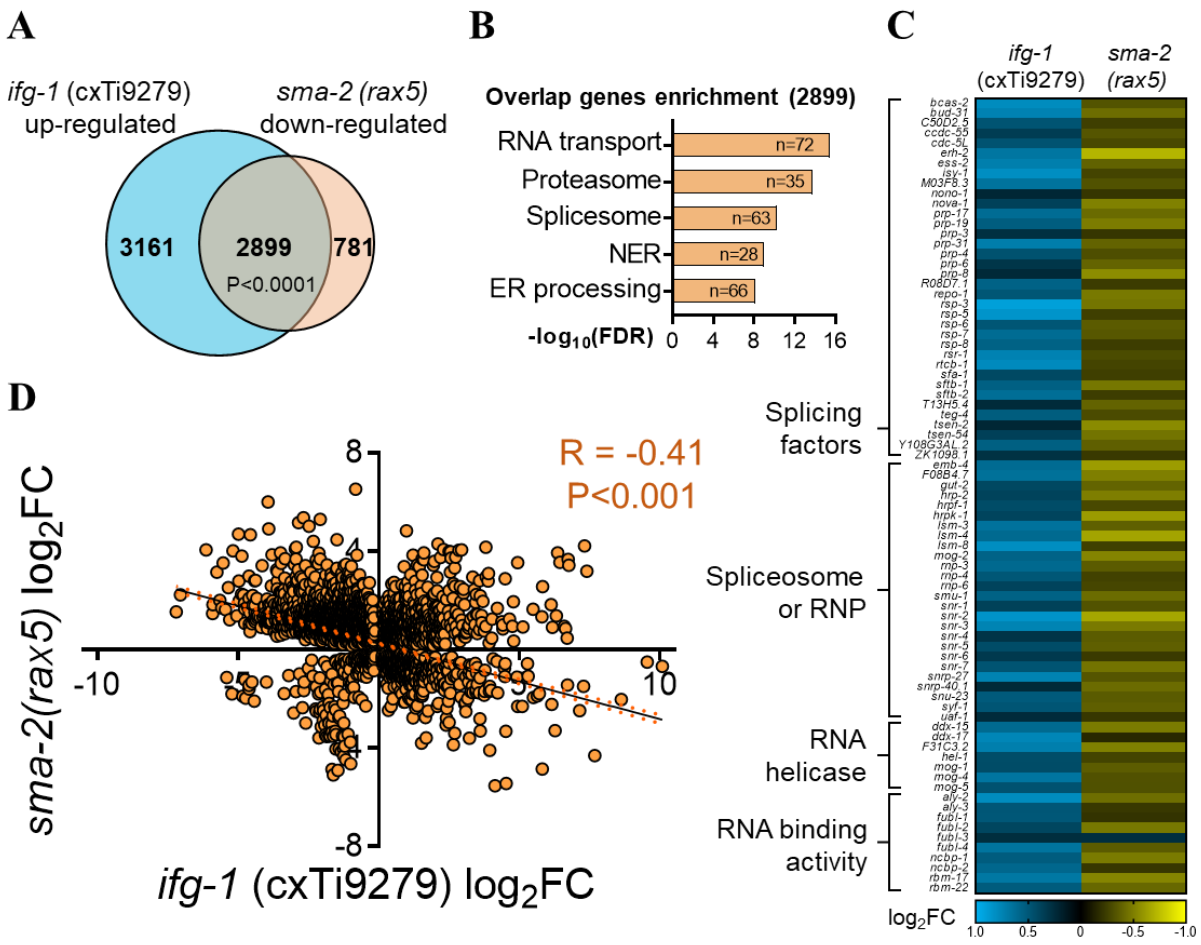


Figure 4.11: SMA-2 regulates *ifg-1* differentially expressed genes

A) Venn diagram showing the overlap between genes upregulated by *ifg-1* mutant and downregulated by *sma-2* mutant. **B)** Enrichment analysis of overlapping 2899 genes in A. NER: Nucleotide excision repair. **C)** Heatmap of RNA splicing regulatory genes upregulated in *ifg-1* and their corresponding expression in *sma-2(rax5)*. **D)** Expression of significantly differentially expressed genes plotted on an XY graph, where X is the gene expression level in *ifg-1* and Y is the gene expression level of the same gene in *sma-2(rax5)*. Each point on the graph represents a gene that is significantly altered in *ifg-1*, and its position on the graph is related to its expression in the *ifg-1* and *sma-2(rax5)* strain.

5. Discussion

5.1: Translation suppression protects RNA splicing fidelity

RNA splicing is catalyzed by the spliceosome via a highly conserved mechanism in which pre-mRNA is excised and joined together to create mature mRNA that functions as the final template to synthesize proteins. The spliceosome includes over a hundred proteins, but five are highly described in literature for their roles in the act of splicing itself. It begins when the snRNP U1 binds to a highly regulated splice site at the 5' end of an intron and cleaves the strand, and binds this excised end to a downstream branch point to create an RNA lariat. Next, the snRNPs U2 and U4/U6 position the excised 5' end with the 3' end, followed by the snRNP U5 excising the 3' end at a splice site. The two excised ends are joined together by the snRNPs present, and the intron lariat is released for degradation. This process is repeated multiple times on a strand of pre-mRNA at predetermined splice sites, resulting in mature mRNA that is transported out of the nucleus for translation. It is essential for RNA splicing to occur rapidly and properly in eukaryotes for efficient translation and cell function; improper RNA splicing, known as RNA mis-spicing, results in inefficient translation that can lead to improper cell function. When RNA mis-splicing occurs, pre-mRNA can be cut at incorrect splice sites, creating an undesired final protein that either has an altered function or no function at all. Under normal conditions, this incorrect product is marked for degradation and removed by the cell via a process called nonsense mediated decay. However, when RNA mis-splicing becomes systemic, these incorrect proteins are not disposed of properly, causing protein aggregates or dangerous circular RNA with no end for degradation to form while the desired protein is not synthesized [10]. This systemic dysregulation is known to lead to disease states such as Alzheimer's disease, certain cancers, early onset aging, as well as other disease states [9]. However, what is not known is the mechanisms behind which RNA mis-splicing occurs, and in extension, how to prevent it. The research performed here aims to take one step towards understanding this unknown.

To begin filling in the blanks in this grey area, a genome-wide RNAi screen was performed using a strain of *C. elegans* containing a GFP reporter to monitor *in vivo* RNA splicing. The basis for this screen is to use RNAi *E. coli* targeting and knocking down one gene at a time to determine genetic factors that protect against cadmium-induced RNA mis-splicing.

After screening the ~20,000 genes in the *C. elegans* genome, a total of 64 validated genes were identified as genes that protected RNA splicing when knocked down (Appendix B). Of note, a total of 42 genes of interest are related to protein synthesis, while the others have various functions (Figure 4.1). This result is quite intriguing as research on the effect of translation suppression on RNA splicing fidelity has not been reported, however, there's a vast amount of data that supports the hypothesis of translation suppression being beneficial for the lifespan of various organisms [50]. While an effect on lifespan cannot be directly applied to an effect on RNA splicing, there is a large amount of literature supporting the hypothesis of RNA mis-splicing leading to premature aging [37], indicating that the previous literature on protein synthesis and aging may be applied to protein synthesis and RNA splicing fidelity. As uncovered here, the suppression of protein synthesis via the knockdown of select translation-related genes is beneficial to the host animal's RNA splicing fidelity under stress. Two possible explanations for this result arise: the first of which is that the suppression of protein synthesis results in an indirect protection of RNA splicing under stress; the second is translation suppression results in a direct protection of RNA splicing under stress. In the first case, the increase in RNA splicing fidelity could be explained by a reallocation of ATP from the energy-consuming process of translation to RNA splicing, where there is more energy available for surveillance against RNA mis-splicing and degrading any erroneous product. The second case implies a direct mechanism between translation suppression and RNA splicing protection under stress, where RNA splicing is protected and enhanced as a method of preventing the damage done to a cell by the loss of proper translation. While the first scenario is plausible due to the high amount of ATP required for translation, the second seems far more likely based on previous literature. The suppression of protein synthesis has positive effects on a whole organism, mainly seen as an increase in lifespan. RNA mis-splicing and AS has increasingly been suggested to be related to aging [37]; my research here may suggest that the increased lifespan linked through the suppression of protein synthesis directly involves RNA splicing regulation.

5.2: Translation suppression extends lifespan and stress resistance

As mentioned previously, the suppression of protein synthesis leads to a highly extended lifespan in various organisms such as *C. elegans* and *Drosophila*. It was first proposed that

translation suppression may allow for an increased lifespan due to a physiological shift of a cell from a normal one to a stressed cell that promotes an increase in maintenance and repair [50], and in 2011 more data surfaced that supported this theory: suppression of genes required for translation resulted in an increase in translation of several stress response genes [52]. Since this discovery, more research has arisen on the topic that supports this theory, although the exact mechanism of how protein synthesis suppression signals to protect lifespan is not well known; all that is known is the cause and effect. In *C. elegans*, genetic regulators of lifespan are well-characterized due to the extreme ease and malleability of the nematode's genetic background [67]. This simple worm is highly regarded as an ideal model for studying aging and effects related to aging not only due to its relatively short lifespan, but also due to the high integration between overall health and lifespan of a worm. Understanding the effect a condition such as chemical exposure or gene knockdown has on the lifespan of a population of *C. elegans* leads to the inherent understanding of that condition's effects on the whole organism.

One gene identified in the RNAi screen discussed in Section 5.1, *ifg-1*, is of particular interest to aging research in *C. elegans*. The only known ortholog of human eIF4G3 present in *C. elegans*, *ifg-1* is a eukaryotic translation initiation factor playing an important role in initiation of translation. The two isoforms of *ifg-1*, a longer p170 form and a shorter p130 form, control cap-dependant or cap-independent mRNA translation respectively, and are expressed differentially in the germline where each type of translation is required [68]. To control translation, an eIF4G and an eIF4E bind together and direct the mRNA strand to a ribosome; as such, they are the first to interact with an mRNA and therefore are the rate-determining factors for the efficiency of translation. The suppression of *ifg-1* is known to cause a direct suppression of translation [51], making it an ideal target for researching the effect of translation suppression on lifespan and RNA splicing. In this research, I employed two different forms of *ifg-1* suppression: RNAi knockdown and a mutated partial loss-of-function strain. When RNAi technology is used in this research, both isoforms of *ifg-1* are targeted unconditionally, resulting in a higher rate of gene knockdown. In the *ifg-1* (cxTi9279) strain, *ifg-1* is mutated via the insertion of a *Mos* transposon that shifts the reading frame of *ifg-1* at exon 5, resulting in the inability to translate the mRNA into the expected protein. The *Mos* transposon insertion results in the decrease of the p170 *ifg-1* isoform and impairs cap-dependent protein translation [59]. Interestingly, the *ifg-1* (cxTi9279) has been shown to be long-lived compared to the wildtype even though the p130 isoform is

relatively unaffected, this suggest that partial loss-of-function to *ifg-1* is sufficient to extend lifespan.

In this thesis, I performed lifespan assays on twelve of the 64 genes identified in the RNAi screen discussed in Section 5.1 in an effort to help understand the relationship between improved RNA splicing fidelity under stress and lifespan. I focused on two categories of genes: first, six genes with previously research literature available that are related to translation; and six more that have little research available that are unrelated to translation. With these data, I aimed to pinpoint specific genes that do or do not confer the long-lived phenotype caused by translation suppression, as well as potentially uncover another class of uncharacterized genes that extends lifespan. The results of this experiment, as seen in Figure 4.2 and Table 4.1, confirm that translation suppression via RNAi (the first six genes listed in each figure) allows for, in general, an increased lifespan. The suppression of *rps-6*, a constituent of the small subunit of the ribosome, did not see the increased lifespan that the other five did; this is in direct conflict with previous literature, where an increase in lifespan was seen after *rps-6* knockdown [50]. This inconsistency may be due to human error, contamination within the RNAi *E. coli* bacterial stock used, or general differences between the methodologies of the two labs. The knockdown of each of the other five genes related to translation all increased lifespan, with lifespans ranging from 4% to 44% longer than worms fed with empty vector RNAi. While this data did not help to pinpoint specific genes that do not convey the expected long-lived phenotype, it assists in confirming the hypothesis of translation suppression increasing lifespan. As for the six novel genes unrelated to translation (the last six genes listed in each figure), lifespan extension was not reliably seen. Each gene had different effects, although the overall effect was null or detrimental to lifespan. Effect on lifespan with respect to worms fed with empty vector RNAi ranged from a 39% decrease to an 8% increase, with one notable outlier resulting in a 29% increase in lifespan (*nhr-61*, Trial 3 in Table 4.2). While research exists to confirm the suppression of other mechanisms in *C. elegans* can help to extend lifespan, it is not surprising that the six genes selected here do not have an effect.

A cadmium sensitivity assay, which is identical to the previously discussed lifespan assay except for the sub-lethal dose cadmium in the agar plates, was also performed with the same twelve genes used in the prior lifespan assay. This was done with the intent to uncover if an

organism uses the selected genes identified in the RNAi screen to protect against RNA splicing stress on the full lifespan of a worm and not just within the 24 hour period used for the RNAi screen. The survival of each population under cadmium stress varied between trials more than seen in the lifespan assay, however an overall trend was seen towards knockdown of the first six genes increased survival under cadmium stress, whereas the knockdown of the other six genes either increased the survival under stress a small amount or had no effect/decreased survival under stress (Figure 4.3). Looking at individual genes, *rps-23* saw the highest increase in survival across all trials, while H12I13.2 (predicted ATP binding activity) saw the most decreased survival (Table 4.2). The knockdown of *ifg-1* resulted in a slight increase in survival during all trials, ranging from 8% to 20%. These data help confirm the theory that translation suppression not only helps to extend lifespan under normal conditions, but also to extend survival under stress. These results together show that the suppression of translation proves beneficial to the lifespan of a worm under both normal and environmental stress based conditions.

5.3: *ifg-1* knockdown protects RNA splicing fidelity

Research on the connection between RNA splicing fidelity and lifespan of an organism has recently seen traction, however this hypothesis is not new; abnormal splicing profiles were first linked to aging in 1977 in the liver of rats [36]. More recently, global genome sequencing of mice showed a definite increase in AS events across the whole organism with aging [39]. Since the mid 2010s, research on the topic has seen a sharp increase and expansion to other organisms such as humans, *Drosophila*, and *C. elegans*. A large amount of literature exists to assist in the hypothesis that RNA mis-splicing and AS events increase as an animal ages, however the mechanism behind this is not well understood. An understanding of this mechanism can assist in curing age-related diseases that currently only have treatments available for their symptoms, such as Alzheimer's disease and cancer. If we can understand the cause of the AS events that lead to early aging and disease onset, as well as how these AS events can be avoided in healthy individuals, then the knowledge can be applied to treating those affected or even preventing the disease state entirely.

In this research, I performed RNA extraction on populations of wild type and *ifg-1* mutant worms with and without cadmium exposure in an effort to understand the whole-transcriptome effects that *ifg-1* produces on worms exposed to cadmium. After RNA extraction, samples were sequenced by Novogene via Illumina platforms, which utilizes "sequencing by synthesis" where a strand is sequenced as it is synthesized. The sequences were returned alongside large amount of data on a wide variety of categories, including data on differential gene analysis and AS events for each sample. Through this, I was able to see a striking similarity in gene expressional changes between two unexpected populations: N2 worms exposed to cadmium and *ifg-1* worms not exposed to cadmium (Figure 4.7C). This response was unexpected – wild type worms under a state of stress appearing to have a similar transcriptome profile to *ifg-1* mutants not under any external stress. While earlier studies have shown that *ifg-1* mutants have increased in expression of select stress response genes, the data here suggest that the effect is far more global with systemic activation of a stress-like transcriptome state. While it is unclear why these similarities exist, I hypothesized that it is due to both populations being in a state of stress. Perhaps the stress induced by cadmium exposure and the stress caused by translation suppression result in similar changes to the transcriptome? Othumpangat *et al.* researched the effect of cadmium exposure on eIF4E in the human HCT15 and PLC/PR/5 cell lines, finding the same result proposed here: exposure to cadmium chloride resulted in ubiquitination and subsequent degradation of the EIF4E protein, causing suppressed translation [69]. As well, Wang et al. researched a similar effect in *C. elegans*, coming to a similar conclusion [70]. These data help to assist in my theory that exposure to cadmium results in a similar transcriptome change to *ifg-1* mutant worms due to the suppressed translation caused by IFE-3 (the *C. elegans* homolog of eIF4E) degradation. In this same analysis, I also found an upregulation of genes that function in RNA splicing in the *ifg-1* mutants (Figure 4.6A). This figure shows the highest levels of upregulation are present in *ifg-1* worms not exposed to cadmium, but there is still a significant upregulation in *ifg-1* worms exposed to cadmium, whereas the N2 worms exposed to cadmium have little to no upregulation at all. This discovery helps support the hypothesis that translation suppression and RNA splicing protection are linked, and even uncovers which specific genes are upregulated to compensate for the RNA splicing disruption caused by cadmium stress.

Next, I analyzed the effects of each of these conditions on the AS profiles produced. Data was available on isoform expression for the five categories of AS events; exon skipping,

mutually exclusive exons, alternative 3’ or 5’ splice site, or intron retention. In the N2 worms exposed to cadmium, there was a highly significant amount of AS events in all five categories, this is consistent with previously published data [14]. What I found in regards to *ifg-1* mutants exposed to cadmium was a statistically significant decrease in AS when compared to wild type worms under identical conditions – by 50%, on average (Figure 4.6B). While these numbers varied per category of AS, there is a clear pattern present: the suppression of translation through an *ifg-1* mutation results in the protection of RNA splicing and therefore the reduction of AS events upon stress exposure. Overall, there was a negative correlation between N2 and *ifg-1* worms both exposed to cadmium, meaning an increase of an isoform in one population was matched by a decrease of that isoform in the other population (Figure 4.6C). There was a significant negative correlative effect seen in the exon skipping category only (Figure 4.6D), likely due to it being the largest category of AS events, allowing for the highest statistical significance. The enormous amount of exon skipping events in relation to the other four types of events combined likely influenced the overall trend to be negative, even though some other categories displayed positive trends (alternative 5’ and 3’ splice site). However, its unknown if one single category of AS events is more influential than another for the health of a cell, I focused on the overall trends seen. With these data, I can suggest that *ifg-1* suppression allows for increased protection of RNA splicing under stress either through the direct reduction of AS events or enabling protection against AS events; likely the latter, as a protection of RNA splicing allows for a healthier mechanism in place to detect and degrade RNA mis-splicing and unwanted AS events.

5.4: Translation suppression induced lifespan extension requires RNA splicing

The data presented so far has shown that a reduction of *ifg-1* results in protected RNA splicing and increased lifespan under normal and stressed conditions. Moving forward, I aimed to understand the necessity of proper RNA splicing fidelity for the lifespan extension provided by *ifg-1* reduction. To do this, lifespans were performed with both wild type worms and the *ifg-1* (cxTi9279) mutant strain on RNAi knockdown *E. coli* targeting a splicing factor of interest. This resulted in a population of *ifg-1* mutants with an additional splicing factor knocked down via RNAi, which allowed me to determine if healthy RNA splicing is required for lifespan extension

under *ifg-1* reduction. The genes that I used for these experiments were all essential for RNA splicing – *snr-1, snr-2, sfa-1, uaf-2,* and *rsp-2.* Two more genes were also selected and studied in this lifespan assay, *hrp-2* and *hrpf-1,* but after sequencing the *E. coli* stocks it was determined that the RNAi plasmid present was targeting an entirely different gene; as such, the data collected will not be assessed here and is instead presented in Appendix D. Each of the five genes studied has confirmed roles in RNA splicing, from the small nuclear riboproteins *snr-1* and *snr-2* to splice site binding or splicing (*uaf*-2 or *sfa-1,* respectively), or general RNA binding activity (*rsp-2*). Overall, the knockdown of splicing factors in an *ifg-1* mutant background resulted in a decrease in lifespan. These data can be seen in Figure 4.8, as well as in Table 4.3. One exception is the knockdown of *rsp-2,* which resulted in no change in lifespan in either the wild type or the *ifg-1* mutant background. The knockdown of either *snr-1* or *snr-2* resulted in a dramatic decrease in lifespan in both backgrounds, with an equal shift seen in each background, meaning the knockdown of the gene in one background did not have a more pronounced effect than the knockdown of the same gene in the other background. And finally, the knockdown of *uaf-2* or *sfa-1* decreased lifespan in the *ifg-1* mutant background but not in the wild-type background. When genes are analyzed individually, the knockdowns that produced the most interesting result were *uaf-2* and *sfa-1* – while wild-type worms were unaffected, *ifg-1* mutants saw a dramatic decrease in lifespan, well below the wild-type lifespan with and without the knockdown present. I hypothesize this to mean that these two genes specifically are required by *ifg-1* for its mechanism of lifespan extension and regulation, but are dispensable in normal aging as they did not affect the wildtype lifespan. Overall, these data indicate for the first time that a functional spliceosome is necessary for the lifespan extension seen under *ifg-1* suppression, and implicates the importance of RNA splicing regulation as a mediator of longevity in long-lived mutants.

Just as before, I performed a cadmium sensitivity assay using the same conditions as the lifespan assay. Wildtype and *ifg-1* mutant strains were used with the same five genes of interest, although this experiment took place on cadmium to induce RNA splicing stress. As seen in the first cadmium sensitivity assay performed, *ifg-1* suppression results in an increased lifespan under stress. I next wanted to test if *ifg-1* requires proper RNA splicing for its stress resistance that allows the animal to live longer under stressful conditions. However, as seen in Figure 4.9 or Table 4.4, there was minimal to no change in cadmium survival of *ifg-1* mutants after RNAi

knockdown of RNA splicing regulating genes compared to EV RNAi. While some minor variations exist in each category, overall the results show that proper RNA splicing is not necessary for the increased cadmium survival induced by *ifg-1* knockdown. A potential explanation is that the RNA splicing pathway is disrupted due cadmium toxicity, as such, further disruption of the spliceosome via RNAi would not necessarily create an additive effect. This is consistent with a previous study that showed disruption of snRNA processing by knocking down the Integrator complex in *C. elegans* shortens lifespan under normal conditions, but exhibit no additive effect under cadmium exposure [14].

An interesting observation in these lifespan and cadmium sensitivity assays compared to the assays discussed in Section 5.2 is a significant difference in lifespan between *ifg-1* knockdown via RNAi or via the partial loss-of-function mutant. The typical lifespan increase seen in *ifg-1* RNAi knockdown populations is about 35% when compared to the wildtype worms literature [51], [52], which agrees with what I saw in the first lifespan (Section 5.2). However, the increase in lifespan discussed in this section was only 10% at best using the *ifg-1* mutant. The explanation for this difference in lifespan extension may be explained by the inherent difference in amount of functional *ifg-1* present in the mutant vs RNAi knockdown. As mentioned before, *ifg-1* has two main isoforms expressed independently in various regions of the worm; p130 and p170. While RNAi knockdown targets both isoforms equally, the cxTi9279 mutant strain used to create the *ifg-1* knockdown strain only targets the p170 isoform. The difference in lifespan between RNAi knockdown and cxTi9279 mutant worms has previously been documented in literature [53], although not to the extent seen in this lifespan assay. The other factor assisting in this dramatic change in lifespan may be due to simple human error; counting worms that die off early in the lifespan as dead worms when they should be censor due to genetic defects can decrease the average lifespan of a population greatly. Additional lifespan assays were performed by other members of the lab where censoring these early malformed deaths helped to increase the average lifespan of the cxTi9279 mutant strain to that of the previously published value (data not shown). The same principles can be applied to the cadmium sensitivity assays performed in this work to explain the relative difference in lifespan seen between RNAi knockdown populations and cxTi9279 mutant populations.

5.5: *ifg-1* signals through the SMA family of proteins

Up until now, I have researched the whole-organism effects of *ifg-1* knockdown and cadmium-induced stress, with an emphasis on how these two conditions affects RNA splicing. Next, I wanted to further explore the genetic mechanism driving *ifg-1* mutant's resistance to RNA splicing disruption. To achieve this, a second RNAi screen was performed, however on a much smaller subset of genes. KH2235 worms were first fed *ifg-1* knockdown RNAi before an additional feeding of *E. coli* RNAi from a small sub-library focused on transcription factors, protein kinases, and protein phosphatases in an effort to identify regulatory genes that work downstream of *ifg-1* to assist it in its RNA splicing protection under stress. Out of the ~900 genes assayed, only three were identified as genes that, when knocked down, inhibited the retention of GFP signal of the *in vivo* RNA splicing reporter under cadmium stress that is typically observed when *ifg-1* is knocked-down. The three genes identified were *sma-2, sma-3,* and *sma-4,* with *sma-3* having the strongest effect (Figure 4.10). First identified in *Drosophila* [71] and homologous to human SMAD proteins, this family of transcription factors respond to signals from the cell wall protein family transforming growth factor-β (TGF-β) to the nucleus, where they signal for the transcription of different genes of interest involvement in cell growth and development [8]. As such, these proteins are necessary for proper health of the cell, and an ideal signalling protein for *ifg-1* to incorporate into the lifespan regulation pathway it controls. Using a previously published RNA-sequencing experiment, I also performed an analysis on a list of upregulated genes identified in my *ifg-1* mutants that overlaps with a list of downregulated genes in *sma-2* loss of function mutants [64] and saw that there was an overwhelmingly large overlap of genes between the two populations. Clustering analysis of the overlapped genes revealed that most of which have functions in RNA metabolism including RNA transport and spliceosome. Based on these results, I propose that the SMA family of transcription factors function downstream of translation suppression induced by *ifg-1* mutant, to potentially enact on the RNA splicing pathway by directly influencing the transcription of genes involved in RNA transport and spliceosome to increase protection of RNA splicing under stress.

5.6: Potential mechanism of lifespan regulation by *ifg-1*

While AS is necessary for genetic diversity in an organism, dysregulation in splicing leads to systemic RNA mis-splicing, causing well known disease states such as cancer and premature aging. One theorized method of treating such disease states is to prevent RNA mis-splicing itself by inducing the existing mechanisms within the cell that protect against RNA mis-splicing. Past literature agrees that *ifg-1*, a translation initiation factor in *C. elegans*, is deeply involved with lifespan regulation in a worm and can extend lifespan when suppressed. Prior to this research, the mechanism of which *ifg-1* regulates lifespan under normal condition and extends lifespan when suppressed was presumably due to the increase in stress response genes. Here I present a potential mechanism of action in which this takes place in *C. elegans*, as seen in Figure 5.1. To start, it has been previously documented in literature that a suppression of *ifg-1* results in a direct decrease in translation [51]. Next, based on the data presented here I present that *ifg-1* signals through the SMA family of transcription factors when suppressed. I next propose that this signal is sent to the nucleus to increase the expression of genes involved in RNA splicing and this functions to protect and maintain RNA splicing fidelity. Since I have seen here that *ifg-1* suppression results in protection of RNA splicing; this was determined by the initial RNAi screen that saw *ifg-1* knockdown worms retain their GFP fluorescence which indicates a healthy RNA splicing mechanism, as well as determined by the RNA-sequencing performed which saw a reduction of AS events by 50% when exposed to cadmium. Overall, my results here propose a mechanism by which *ifg-1* suppression leads to an extended lifespan in *C. elegans* by directly protecting RNA splicing fidelity to ensure the homeostasis of RNA metabolism under stress.

5.7: Final Conclusions

Neurodegenerative diseases are characterized by the buildup of protein aggregates in a cell, however the cause of the mass amount of improper protein translation and degradation is poorly understood. Here, I detail one potential mechanism of how this happens, specifically through the *ifg-1* gene and splicing dysregulation in *C. elegans*. Understanding the cause of neurodegenerative diseases through research like this can allow for a greater understanding of how to treat or even cure neurodegenerative diseases before they start.

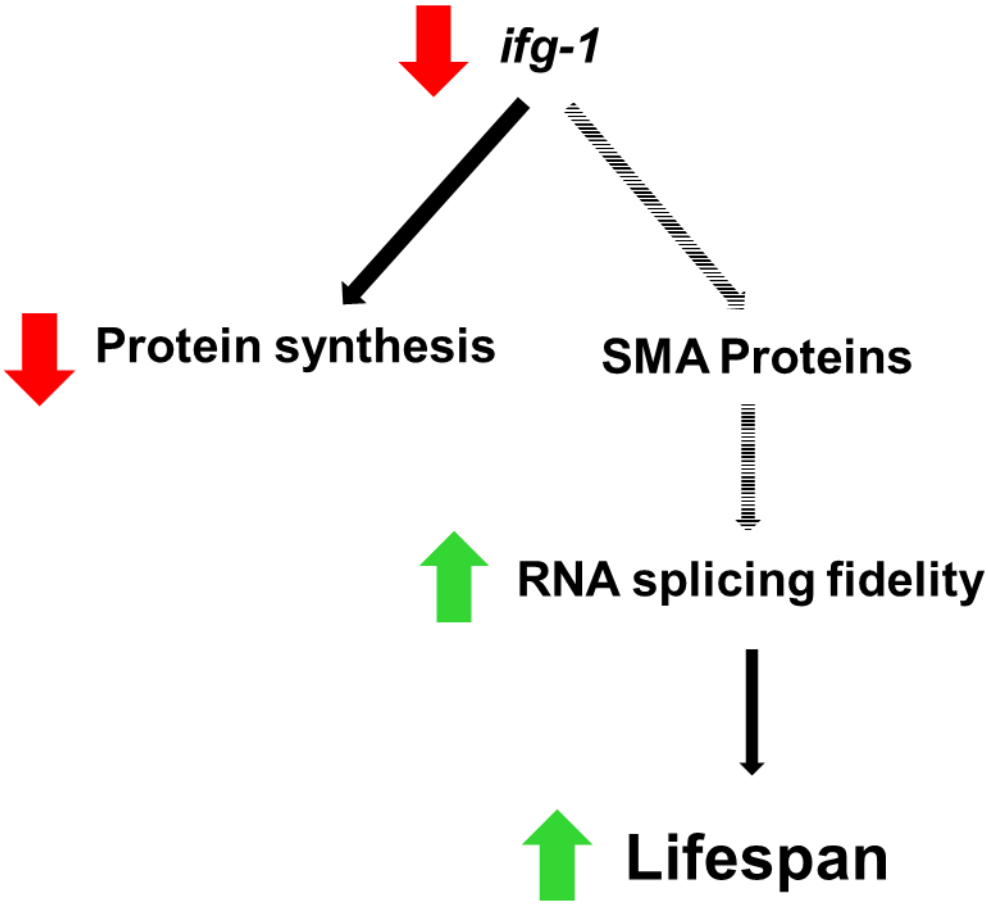


Figure 5.1: Proposed mechanism of action of *ifg-1*'s lifespan extension under stress

In this mechanism, suppression of *ifg-1* leads to a suppression of translation, while also signalling through the SMA family of transcription factors to increase RNA splicing fidelity and improve lifespan. Red arrow: decrease; green arrow: increase; solid black arrow: pathway of the mechanism; dashed black arrow: signalling pathway.

5.8: Strengths and Weaknesses

The use of the model organism *C. elegans* allowed for a wide variety of testing, but limitations are present in this research. *C. elegans* has a rapid lifespan, reaching reproductive maturity in only two days; this allowed me to perform whole-lifespan experiments very quickly. As well, the small size and asexual reproduction allowed for large populations of genetically identical worms for large sample sizes, reaching the hundreds of samples per replicate. However, as with the use of any model organism, extrapolation beyond the organism is complicated. *C. elegans* has a high homology with humans, but further testing is needed before any conclusions can be made on human health based on the data presented here.

5.9: Future work

With this research, I have only scratched the surface on how *ifg-1* regulates lifespan in *C. elegans*. More research needs to be done before a clear understanding of each step of the mechanism can be understood. One experiment that could be done is another RNA-seq of *ifg-1* mutants with RNAi knockdown of the *sma* genes identified in Section 5.5 to understand the whole-organism effects this knockdown has on AS under stress. Looking more broadly at what needs to be done, the specific genes used between each step of this simplified mechanism need to be better understood in *C. elegans*, then in higher organisms with the goal of applying this information to humans. One method of doing this is via knockdown of different protein synthesis related genes, and not just *ifg-1*. This can help confirm that the effects on RNA splicing and lifespan seen here are indeed related to protein synthesis overall and not just specific to *ifg-1* knockdown. Similarly, by performing the experiments done here with stressors other than cadmium such as heat shock, the data seen here can be confirmed to be a result of general RNA splicing disruption and not specifically to an effect of cadmium toxicity. Once this area of research is better understood, we can apply this knowledge to areas of medicine such as devising treatments for disease states brought about by systemic RNA splicing dysregulation.

6. References

[1] T. W. Nilsen, “The spliceosome: the most complex macromolecular machine in the cell?,” *Bioessays*, vol. 25, no. 12, pp. 1147–1149, Dec. 2003, doi: 10.1002/bies.10394.

[2] J. Sambrook, “Adenovirus amazes at Cold Spring Harbor,” *Nature*, vol. 268, no. 5616, pp. 102–104, Jul. 1977, doi: 10.1038/268101a0.

[3] W. Gilbert, “Why genes in pieces?,” *Nature*, vol. 271, no. 5645, p. 501, Feb. 1978, doi: 10.1038/271501a0.

[4] E. A. Ponomarenko *et al.*, “The size of the human proteome: The width and depth,” *International Journal of Analytical Chemistry*, vol. 2016, pp. 1–6, 2016, doi: 10.1155/2016/7436849.

[5] P. J. Thul *et al.*, “A subcellular map of the human proteome,” *Science*, vol. 356, no. 6340, p. eaal3321, May 2017, doi: 10.1126/science.aal3321.

[6] Q. Pan, O. Shai, L. J. Lee, B. J. Frey, and B. J. Blencowe, “Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing,” *Nat Genet*, vol. 40, no. 12, pp. 1413–1415, Dec. 2008, doi: 10.1038/ng.259.

[7] Y. Liu, P. Li, L. Fan, and M. Wu, “The nuclear transportation routes of membrane-bound transcription factors,” *Cell Commun Signal*, vol. 16, no. 1, p. 12, Dec. 2018, doi: 10.1186/s12964-018-0224-3.

[8] C. Savage, P. Das, A. L. Finelli, S. R. Townsend, C.-Y. Sunt, and S. E. Bairdt, “*Caenorhabditis elegan*s genes *sma-2, sma-3,* and *sma-4* define a conserved family of transforming growth factor 3 pathway components,” *Developmental Biology*, p. 5, 1996.

[9] M. M. Scotti and M. S. Swanson, “RNA mis-splicing in disease,” *Nat Rev Genet*, vol. 17, no. 1, pp. 19–32, Jan. 2016, doi: 10.1038/nrg.2015.3.

[10] C. Cocquerelle, B. Mascrez, D. Hétuin, and B. Bailleul, “Mis-splicing yields circular RNA molecules,” *The FASEB Journal*, vol. 7, no. 1, pp. 155–160, Jan. 1993, doi: 10.1096/fasebj.7.1.7678559.

[11] G. Dujardin *et al.*, “How slow RNA polymerase II elongation favors alternative exon skipping,” *Molecular Cell*, vol. 54, no. 4, pp. 683–690, May 2014, doi: 10.1016/j.molcel.2014.03.044.

[12] G. Biamonti and J. F. Caceres, “Cellular stress and RNA splicing,” *Trends in Biochemical Sciences*, vol. 34, no. 3, pp. 146–153, Mar. 2009, doi: 10.1016/j.tibs.2008.11.004.

[13] J. A. Pleiss, G. B. Whitworth, M. Bergkessel, and C. Guthrie, “Rapid, transcript-specific changes in splicing in response to environmental stress,” *Molecular Cell*, vol. 27, no. 6, pp. 928–937, Sep. 2007, doi: 10.1016/j.molcel.2007.07.018.

[14] C.-W. Wu, K. Wimberly, A. Pietras, W. Dodd, M. B. Atlas, and K. P. Choe, “RNA processing errors triggered by cadmium and integrator complex disruption are signals for environmental stress,” *BMC Biology*, vol. 17, no. 1, Dec. 2019, doi: 10.1186/s12915-019-0675-z.

[15] W. E. Balch, R. I. Morimoto, A. Dillin, and J. W. Kelly, “Adapting proteostasis for disease intervention,” *Science*, vol. 319, no. 5865, pp. 916–919, Feb. 2008, doi: 10.1126/science.1141448.

[16] B. Bai *et al.*, “U1 small nuclear ribonucleoprotein complex and RNA splicing alterations in Alzheimer’s disease,” *Proceedings of the National Academy of Sciences*, vol. 110, no. 41, pp. 16562–16567, Oct. 2013, doi: 10.1073/pnas.1310249110.

[17] I. Diner *et al.*, “Aggregation properties of the small nuclear ribonucleoprotein U1-70K in Alzheimer Disease,” *Journal of Biological Chemistry*, vol. 289, no. 51, pp. 35296–35313, Dec. 2014, doi: 10.1074/jbc.M114.562959.

[18] C. M. Hales *et al.*, “U1 small nuclear ribonucleoproteins (snRNPs) aggregate in Alzheimer’s disease due to autosomal dominant genetic mutations and trisomy 21,” *Mol Neurodegeneration*, vol. 9, no. 1, p. 15, 2014, doi: 10.1186/1750-1326-9-15.

[19] L. Maroteaux and R. H. Scheller, “The rat brain synucleins; family of proteins transiently associated with neuronal membrane,” *Molecular Brain Research*, vol. 11, no. 3–4, pp. 335–343, Oct. 1991, doi: 10.1016/0169-328X(91)90043-W.

[20] I. D’Souza *et al.*, “Missense and silent tau gene mutations cause frontotemporal dementia with parkinsonism-chromosome 17 type, by affecting multiple alternative RNA splicing regulatory elements,” *Proceedings of the National Academy of Sciences*, vol. 96, no. 10, pp. 5598–5603, May 1999, doi: 10.1073/pnas.96.10.5598.

[21] L. A. Shehadeh *et al.*, “SRRM2, a potential blood biomarker revealing high alternative splicing in Parkinson’s Disease,” *PLoS ONE*, vol. 5, no. 2, p. e9104, Feb. 2010, doi: 10.1371/journal.pone.0009104.

[22] O. Korvatska *et al.*, “Altered splicing of ATP6AP2 causes X-linked parkinsonism with spasticity (XPDS),” *Human Molecular Genetics*, vol. 22, no. 16, pp. 3259–3268, Aug. 2013, doi: 10.1093/hmg/ddt180.

[23] K. Yoshida *et al.*, “Frequent pathway mutations of splicing machinery in myelodysplasia,” *Nature*, vol. 478, no. 7367, pp. 64–69, Oct. 2011, doi: 10.1038/nature10496.

[24] H. Jung, K. S. Lee, and J. K. Choi, “Comprehensive characterisation of intronic mis-splicing mutations in human cancers,” *Oncogene*, vol. 40, no. 7, pp. 1347–1361, Feb. 2021, doi: 10.1038/s41388-020-01614-3.

[25] E. El Marabti and I. Younis, “The cancer spliceome: Reprograming of alternative splicing in cancer,” *Front. Mol. Biosci.*, vol. 5, p. 80, Sep. 2018, doi: 10.3389/fmolb.2018.00080.

[26] P. I. Poulikakos *et al.*, “RAF inhibitor resistance is mediated by dimerization of aberrantly spliced BRAF(V600E),” *Nature*, vol. 480, no. 7377, pp. 387–390, Dec. 2011, doi: 10.1038/nature10662.

[27] A. S. Adler *et al.*, “An integrative analysis of colon cancer identifies an essential function for PRPF6 in tumor growth,” *Genes & Development*, vol. 28, no. 10, pp. 1068–1084, May 2014, doi: 10.1101/gad.237206.113.

[28] S. Bonnal, L. Vigevani, and J. Valcárcel, “The spliceosome as a target of novel antitumour drugs,” *Nat Rev Drug Discov*, vol. 11, no. 11, pp. 847–859, Nov. 2012, doi: 10.1038/nrd3823.

[29] T. Y.-T. Hsu *et al.*, “The spliceosome is a therapeutic vulnerability in MYC-driven cancer,” *Nature*, vol. 525, no. 7569, pp. 384–388, Sep. 2015, doi: 10.1038/nature14985.

[30] R. Pearl, “The rate of living.” University of London Press, 1928.

[31] J. Bjorksten, “The crosslinkage theory of aging,” *Journal of the American Geriatrics Society*, vol. 16, no. 4, pp. 408–427, Apr. 1968, doi: 10.1111/j.1532-5415.1968.tb02821.x.

[32] R. Gerschman, D. L. Gilbert, S. W. Nye, P. Dwyer, and W. O. Fenn, “Oxygen poisoning and X-irradiation: A mechanism in common,” *Science*, vol. 119, no. 3097, pp. 623–626, May 1954, doi: 10.1126/science.119.3097.623.

[33] D. Harman, “Aging: A theory based on free radical and radiation chemistry,” *Journal of Gerontology*, vol. 11, no. 3, pp. 298–300, Jul. 1956, doi: 10.1093/geronj/11.3.298.

[34] H. L. Gensler and H. Bernstein, “DNA damage as the primary cause of aging,” *The Quarterly Review of Biology*, vol. 56, no. 3, pp. 279–303, Sep. 1981, doi: 10.1086/412317.

[35] K. K. Steffen and A. Dillin, “A ribosomal perspective on proteostasis and aging,” *Cell Metabolism*, vol. 23, no. 6, pp. 1004–1012, Jun. 2016, doi: 10.1016/j.cmet.2016.05.013.

[36] A. Yannarell, D. E. Schumm, and T. E. Webb, “Age-dependence of nuclear RNA processing,” *Mechanisms of Ageing and Development*, vol. 6, pp. 259–264, Jan. 1977, doi: 10.1016/0047-6374(77)90026-4.

[37] M. Bhadra, P. Howell, S. Dutta, C. Heintz, and W. B. Mair, “Alternative splicing in aging and longevity,” *Hum Genet*, vol. 139, no. 3, pp. 357–369, Mar. 2020, doi: 10.1007/s00439-019-02094-6.

[38] H. Gruner, M. Cortés-López, D. A. Cooper, M. Bauer, and P. Miura, “CircRNA accumulation in the aging mouse brain,” *Sci Rep*, vol. 6, no. 1, p. 38907, Dec. 2016, doi: 10.1038/srep38907.

[39] S. A. Rodríguez *et al.*, “Global genome splicing analysis reveals an increased number of alternatively spliced genes with aging,” *Aging Cell*, vol. 15, no. 2, pp. 267–278, Apr. 2016, doi: 10.1111/acel.12433.

[40] The *C. elegans* Sequencing Consortium, “Genome sequence of the nematode *C. elegans*: A platform for investigating biology,” *Science*, vol. 282, no. 5396, pp. 2012–2018, Dec. 1998, doi: 10.1126/science.282.5396.2012.

[41] A. K. Corsi, “A transparent window into biology: A primer on *Caenorhabditis elegans*,” *WormBook*, pp. 1–31, Jun. 2015, doi: 10.1895/wormbook.1.177.1.

[42] S. Brenner, “The genetics of *Caenorhabditis elegans*,” *Genetics*, vol. 77, no. 1, pp. 71–94, May 1974.

[43] C.-H. Lai, “Identification of novel human genes evolutionarily conserved in *Caenorhabditis elegans* by comparative proteomics,” *Genome Research*, vol. 10, no. 5, pp. 703–713, May 2000, doi: 10.1101/gr.10.5.703.

[44] M. R. Klass, “Aging in the nematode *Caenorhabditis elegans*: Major biological and environmental factors influencing life span,” *Mechanisms of Ageing and Development*, vol. 6, pp. 413–429, Jan. 1977, doi: 10.1016/0047-6374(77)90043-4.

[45] D. B. Friedman and T. E. Johnson, “A mutation in the *age-1* gene in *Caenorhabditis elegans* lengthens life and reduces hermaphrodite fertility.,” *Genetics*, vol. 118, no. 1, pp. 75–86, Jan. 1988, doi: 10.1093/genetics/118.1.75.

[46] B. J. Morris, D. C. Willcox, T. A. Donlon, and B. J. Willcox, "*FOXO3:* A major gene for human longevity - A mini-review," *Gerontology*, vol. 61, no. 6, pp. 515–525, 2015, doi: 10.1159/000375235.

[47] M. Kaeberlein, "Regulation of yeast replicative life span by TOR and Sch9 in response to nutrients," *Science*, vol. 310, no. 5751, pp. 1193–1196, Nov. 2005, doi: 10.1126/science.1115535.

[48] T. L. Kaeberlein *et al.*, "Lifespan extension in *Caenorhabditis elegans* by complete removal of food," *Aging Cell*, vol. 5, no. 6, pp. 487–494, Dec. 2006, doi: 10.1111/j.1474-9726.2006.00238.x.

[49] B. P. Braeckman and J. R. Vanfleteren, "Genetic control of longevity in *C. elegans*," *Experimental Gerontology*, vol. 42, no. 1–2, pp. 90–98, Jan. 2007, doi: 10.1016/j.exger.2006.04.010.

[50] M. Hansen, S. Taubert, D. Crawford, N. Libina, S.-J. Lee, and C. Kenyon, "Lifespan extension by conditions that inhibit translation in *Caenorhabditis elegans*," *Aging Cell*, vol. 6, no. 1, pp. 95–110, Feb. 2007, doi: 10.1111/j.1474-9726.2006.00267.x.

[51] K. Z. Pan *et al.*, "Inhibition of mRNA translation extends lifespan in *Caenorhabditis elegans*," *Aging Cell*, vol. 6, no. 1, pp. 111–119, Feb. 2007, doi: 10.1111/j.1474-9726.2006.00266.x.

[52] A. N. Rogers *et al.*, "Life span extension via eIF4G inhibition is mediated by posttranscriptional remodeling of stress response gene expression in *C. elegans*," *Cell Metabolism*, vol. 14, no. 1, pp. 55–66, Jul. 2011, doi: 10.1016/j.cmet.2011.05.010.

[53] A. C. Howard, J. Rollins, S. Snow, S. Castor, and A. N. Rogers, "Reducing translation through eIF4G/IFG-1 improves survival under ER stress that depends on heat shock factor HSF-1 in *Caenorhabditis elegans*," *Aging Cell*, vol. 15, no. 6, pp. 1027–1038, Dec. 2016, doi: 10.1111/acel.12516.

[54] J.-F. Rual, "Toward improving *Caenorhabditis elegans* phenome mapping with an ORFeome-based RNAi library," *Genome Research*, vol. 14, no. 10b, pp. 2162–2168, Oct. 2004, doi: 10.1101/gr.2505604.

[55] R. Kamath, "Genome-wide RNAi screening in *Caenorhabditis elegans*," *Methods*, vol. 30, no. 4, pp. 313–321, Aug. 2003, doi: 10.1016/S1046-2023(03)00050-1.

[56] T. Stiernagle, “Maintenance of *C. elegans*,” *WormBook*, 2006, doi: 10.1895/wormbook.1.101.1.

[57] B. D. Williams, B. Schrank, C. Huynh, R. Shownkeen, and R. H. Waterston, “A genetic mapping system in *Caenorhabditis elegans* based on polymorphic sequence-tagged sites,” *Genetics*, vol. 131, no. 3, pp. 609–624, Jul. 1992, doi: 10.1093/genetics/131.3.609.

[58] C. Heintz *et al.*, “Splicing factor 1 modulates dietary restriction and TORC1 pathway longevity in *C. elegans*,” *Nature*, vol. 541, no. 7635, pp. 102–106, Jan. 2017, doi: 10.1038/nature20789.

[59] J. K. Morrison, A. J. Friday, M. A. Henderson, E. Hao, and B. D. Keiper, “Induction of cap-independent BiP (*hsp-3*) and Bcl-2 (*ced-9*) translation in response to eIF4G (IFG-1) depletion in *C. elegans*,” *Translation*, vol. 2, no. 1, p. e28935, Jan. 2014, doi: 10.4161/trla.28935.

[60] J. Ahringer, “Reverse genetics,” *WormBook*, 2006, doi: 10.1895/wormbook.1.47.1.

[61] F. R. G. Amrit, R. Ratnappan, S. A. Keith, and A. Ghazi, “The *C. elegans* lifespan assay toolkit,” *Methods*, vol. 68, no. 3, pp. 465–475, Aug. 2014, doi: 10.1016/j.ymeth.2014.04.002.

[62] S. K. Han *et al.*, “OASIS 2: online application for survival analysis 2 with features for the analysis of maximal lifespan and healthspan in aging research,” *Oncotarget*, vol. 7, no. 35, pp. 56147–56152, Aug. 2016, doi: 10.18632/oncotarget.11269.

[63] A. Fire, S. Xu, M. K. Montgomery, S. A. Kostas, S. E. Driver, and C. C. Mello, “Potent and specific genetic interference by double-stranded RNA in,” vol. 391, p. 6, 1998.

[64] Y. Yu, A. S. Mutlu, H. Liu, and M. C. Wang, “High-throughput screens using photo-highlighting discover BMP signaling in mitochondrial lipid oxidation,” *Nat Commun*, vol. 8, no. 1, p. 865, Dec. 2017, doi: 10.1038/s41467-017-00944-3.

[65] D. W. Huang, B. T. Sherman, and R. A. Lempicki, “Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists,” *Nucleic Acids Research*, vol. 37, no. 1, pp. 1–13, Jan. 2009, doi: 10.1093/nar/gkn923.

[66] D. W. Huang, B. T. Sherman, and R. A. Lempicki, “Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources,” *Nat Protoc*, vol. 4, no. 1, pp. 44–57, Jan. 2009, doi: 10.1038/nprot.2008.211.

[67] A. S. Anisimova, A. I. Alexandrov, N. E. Makarova, V. N. Gladyshev, and S. E. Dmitriev, "Protein synthesis and quality control in aging," *aging*, vol. 10, no. 12, pp. 4269–4288, Dec. 2018, doi: 10.18632/aging.101721.

[68] H. P. Huggins *et al.*, "Distinct roles of two eIF4E isoforms in the germline of *Caenorhabditis elegans*," *Journal of Cell Science*, p. 25, 2020.

[69] S. Othumpangat, M. Kashon, and P. Joseph, "Eukaryotic Translation Initiation Factor 4E Is a cellular target for toxicity and death due to exposure to cadmium chloride," *Journal of Biological Chemistry*, vol. 280, no. 26, pp. 25162–25169, Jul. 2005, doi: 10.1074/jbc.M414303200.

[70] J. Wang, S. Robida-Stubbs, J. M. A. Tullet, J.-F. Rual, M. Vidal, and T. K. Blackwell, "RNAi screening implicates a SKN-1–dependent transcriptional response in stress resistance and longevity deriving from translation inhibition," *PLoS Genet*, vol. 6, no. 8, p. e1001048, Aug. 2010, doi: 10.1371/journal.pgen.1001048.

[71] L. A. Raftery and D. J. Sutherland, "TGF-β family signal transduction in *Drosophila* development: From *mad* to smads," p. 18, 1999.

[72] D. R. Bentley *et al.*, "Accurate whole human genome sequencing using reversible terminator chemistry," *Nature*, vol. 456, no. 7218, pp. 53–59, Nov. 2008, doi: 10.1038/nature07517.

7. Appendices

Appendix A: RNA-seq as performed by Novogene Co., Ltd.

Novogene performed sequencing on twelve samples I provided them: three N2 control samples (NC); three N2 cadmium samples (NCd); three *ifg-1* (cxTi9279) knockdown control samples (IC); and three *ifg-1* (cxTi9279) knockdown cadmium samples (ICd).

Sequencing took place via Illumina platforms, a highly adopted next-generation sequencing method based off of the sequencing by synthesis (SBS) mechanism [72], which adds a fluorescent terminator tag unique for that dNTP each time a dNTP is added to the growing strand. Once added, the strand is imaged, and the tag is immediately removed to allow for continued transcription of the strand. This results in the reading of a complete sequence of a strand as it is being made, allowing for less error-prone long-read sequences.

Novogene sequenced our samples with the following tags: Batch ID: X202SC20113916-Z01-F001; Species and Version: ensembl_caenorhabditis_elegans_wbcel235_gca_000002985_3; Report Time: 2021-01-02. A simplified workflow of the method of sequencing can be seen in Figure A.1, and a simplified workflow of the bioinformatics analysis that took place after can be seen in Figure A.2. As well, a summary of data quality produced is available in Table A.1.

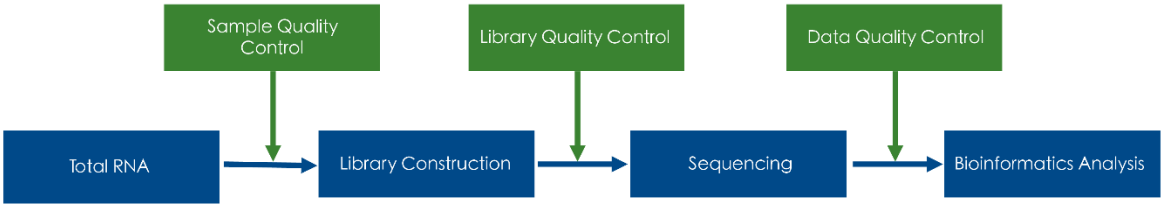


Figure A.1: Workflow of RNA-seq as performed by Novogene

Samples are processed from left to right, and each green box indicates a quality control check that takes place before moving onto the next step.

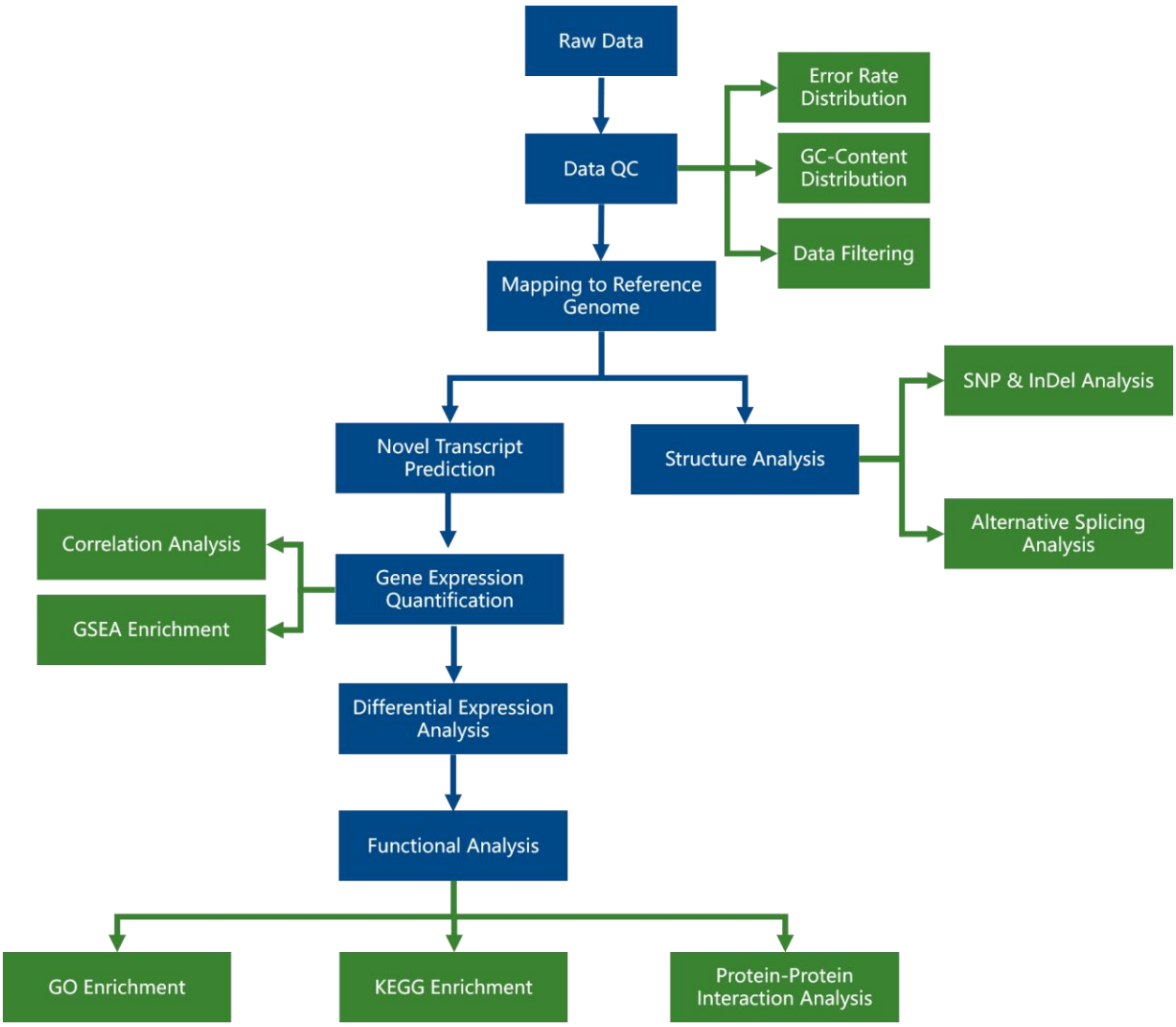


Figure A.2: Workflow of bioinformatics analysis for mRNA sequencing as performed by Novogene. Analysis takes place following the flowchart from top to bottom.

Table A.1: Data quality summary of samples sequenced by Novogene.

Sample name	Raw reads	Clean reads	Raw bases	Clean bases	Error rate (%)	Q20 (%)	Q30 (%)	GC content (%)
NC1	42142432	40614704	12.6G	12.2G	0.02	98.13	94.40	45.36
NC2	47736213	46149283	14.3G	13.8G	0.02	98.03	94.19	45.72
NC3	49310472	47561060	14.8G	14.3G	0.02	98.03	94.19	45.71
NCd1	42734676	41339009	12.8G	12.4G	0.03	97.94	93.96	45.70
NCd2	41288519	39706383	12.4G	11.9G	0.02	98.10	94.34	45.72
NCd3	39867049	38763389	12.0G	11.6G	0.02	98.11	94.37	45.64
IC1	57811365	56171462	17.3G	16.9G	0.02	98.41	95.11	46.62
IC2	42235273	40854204	12.7G	12.3G	0.02	98.24	94.66	46.60
IC3	48022559	46600293	14.4G	14.0G	0.02	98.18	94.54	46.60
ICd1	44325210	43076855	13.3G	12.9G	0.02	98.13	94.41	46.18
ICd2	43937385	42484922	13.2G	12.7G	0.02	98.16	94.46	46.18
ICd3	46136020	44397026	13.8G	13.3G	0.02	98.12	94.38	46.02

Appendix B: RNAi Screen Raw Data

Table B.1: Recorded data from the RNAi screen described in Section 3.1.

Gene Name	**Protein Function**	**Size**	**Fluorescence Level**
acbe-1	ABC transporter, class E	Medium	Strong
aars-23	Alanyl tRNA Synthetase	Small	Strong
H12I13.2	ATP Binding Activity	Medium	Strong
csp-2	CaSPase	Small	Strong
szy-4	Centrosome Duplication	Small	Strong
F46F11.1	Diphosphoinositol-pentakisphosphate kinase	Small	Weak
eef-1A.2	Elongation Factor	Small	Strong
spt-16	FACT complex subunit spt-16	N/A	Weak
fbxa-203	F-box A protein	Small	Weak
B0250.7	Function Unknown	Small	Strong
H12I13.3	Function Unknown	Large	Weak
ekl-7	Germline, Intestine Activity	N/A	Strong
tbcd-1	GTPase activator for Microtubule Organization	Small	Strong
hsf-2	Head Neuron Expression	N/A	Weak
hsp-2	Heat Shock Protein	Small	Weak
T01C3.11	Hypodermis/Intestine	Small	Medium
ifg-1	Initiation Factor 4G (eIF4G) family	Small	Medium
iars-1	Isoleucyl tRNA Synthetase	Large	Strong
nhr-61	Nuclear Hormone Receptor family	Small	Strong
wago-11	Nucleic Acid Binding	Small	Weak
alg-4	Nucleic Acid Binding	Small	Medium
F36A2.3	Oxidoreductace Activity	N/A	Strong
fars-3	phenylalanyl (F) tRNA Synthetase	N/A	Strong
Y48B6A.13	Predicted Diphosphomevalonate Decarboxylase	Small	Medium
BMS1	Predicted GTP Binding Activity	Large	Strong
K08F8.5	Predicted Motility	Small	Strong
K10B4.1	Predicted Peroxidase Activity	Small	Medium
lect-2	Predicted Structural	Large	Strong
eif-2 γ	Predicted tRNA Binding Activity	Small	Strong
M03F8.3	Pre-mRNA Splicing Factor	Small	Strong
pbs-1	Proteasome Beta Subunit	Small	Medium

rla-0	Replication Protein A homolog	Small	Medium
rpl-24.2	Ribosomal Protein, Large subunit	N/A	Strong
rpl-19	Ribosomal Protein, Large subunit	Small	Medium
rpl-36.A	Ribosomal Protein, Large subunit	Small	Strong
rpl-21	Ribosomal Protein, Large subunit	Small	Strong
rpl-22	Ribosomal Protein, Large subunit	Small	Medium
rpl-27	Ribosomal Protein, Large subunit	Small	Medium
rpl-3	Ribosomal Protein, Large subunit	Small	Strong
rpl-26	Ribosomal Protein, Large subunit	Medium	Medium
rpl-25.2	Ribosomal Protein, Large subunit	Large	Medium
rpl-9	Ribosomal Protein, Large subunit	Small	Weak
rps-11.1	Ribosomal Protein, Large subunit	Small	Weak
rpl-30	Ribosomal Protein, Large subunit	Small	Strong
rpl-7A	Ribosomal Protein, Large subunit	Small	Strong
rpl-17	Ribosomal Protein, Large subunit	Small	Medium
rps-0	Ribosomal Protein, Small subunit	Small	Strong
rps-13	Ribosomal Protein, Small subunit	Small	Strong
rps-3	Ribosomal Protein, Small subunit	Small	Strong
rps-30	Ribosomal Protein, Small subunit	Small	Strong
rps-23	Ribosomal Protein, Small subunit	Small	Medium
rps-22	Ribosomal Protein, Small subunit	Small	Weak
rps-12	Ribosomal Protein, Small subunit	Small	Strong
rps-27	Ribosomal Protein, Small subunit	Small	Medium
rps-16	Ribosomal Protein, Small subunit	Small	Strong
rps-5	Ribosomal Protein, Small subunit	Small	Medium
rps-19	Ribosomal Protein, Small subunit	N/A	Strong
rps-20	Ribosomal Protein, Small subunit	Small	Strong
rps-28	Ribosomal Protein, Small subunit	Small	Strong
rps-6	Ribosomal Protein, Small subunit	Small	Strong
eif-1	Ribosomal, Small Subunit Binding	Medium	Weak
T19A6.4	Transmembrane Transport	N/A	Strong
F47B3.6	Tyrosine Phosphatase Activity	Large	Strong

Appendix C: Statistical analysis of qPCR results

Table C.1: Statistical analysis of qPCR results. Two-way ANOVA performed by GraphPad Prism version 8.2.1 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com on the qPCR discussed in Results Section 4.4.

Tukey's multiple comparisons test	**Mean Diff.**	**95.00% CI of diff.**	**Significance**	**Adjusted P Value**
cdr-1				
EV (RNAi):Control vs. EV (RNAi):Cadmium	-327.6	-437.1 to -218.1	****	<0.0001
EV (RNAi):Control vs. *ifg-1* (RNAi):Control	-1.553	-111.0 to 107.9	ns	>0.9999
EV (RNAi):Control vs. *ifg-1* (RNAi):Cadmium	-1.928	-111.4 to 107.6	ns	>0.9999
EV (RNAi):Cadmium vs. *ifg-1* (RNAi):Control	326.1	216.6 to 435.6	****	<0.0001
EV (RNAi):Cadmium vs. *ifg-1* (RNAi):Cadmium	325.7	216.2 to 435.2	****	<0.0001
ifg-1 (RNAi):Control vs. *ifg-1* (RNAi):Cadmium	-0.375	-109.9 to 109.1	ns	>0.9999
numr-1				
EV (RNAi):Control vs. EV (RNAi):Cadmium	-180.3	-308.9 to -51.74	**	0.0062
EV (RNAi):Control vs. *ifg-1* (RNAi):Control	-0.1375	-128.7 to 128.5	ns	>0.9999
EV (RNAi):Control vs. *ifg-1* (RNAi):Cadmium	-122.9	-251.5 to 5.732	ns	0.0628
EV (RNAi):Cadmium vs. *ifg-1* (RNAi):Control	180.2	51.60 to 308.8	**	0.0063
EV (RNAi):Cadmium vs. *ifg-1* (RNAi):Cadmium	57.47	-71.13 to 186.1	ns	0.5645
ifg-1 (RNAi):Control vs. *ifg-1* (RNAi):Cadmium	-122.7	-251.3 to 5.869	ns	0.0631
hsp-16.49				
EV (RNAi):Control vs. EV (RNAi):Cadmium	0.4725	-0.3167 to 1.262	ns	0.3297

EV (RNAi):Control vs. *ifg-1* (RNAi):Control	0.835	0.04578 to 1.624	*	0.0371
EV (RNAi):Control vs. *ifg-1* (RNAi):Cadmium	0.6725	-0.1167 to 1.462	ns	0.1052
EV (RNAi):Cadmium vs. *ifg-1* (RNAi):Control	0.3625	-0.4267 to 1.152	ns	0.5433
EV (RNAi):Cadmium vs. *ifg-1* (RNAi):Cadmium	0.2	-0.5892 to 0.9892	ns	0.8739
ifg-1 (RNAi):Control vs. *ifg-1* (RNAi):Cadmium	-0.1625	-0.9517 to 0.6267	ns	0.9265

Appendix D: Lifespan data for RNAi originally thought to be targeting *hrp-2* and *hrpf-1*

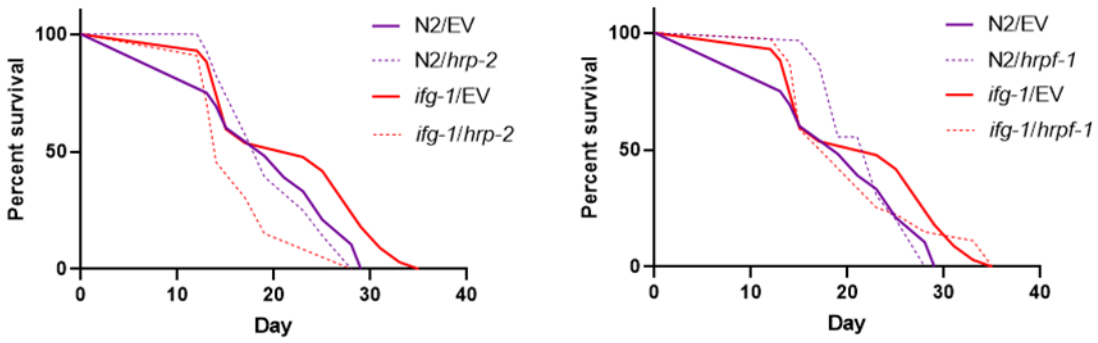


Figure D.1: *ifg-1* mutant background lifespan assay data for RNAi originally thought to be targeting *hrp-2* and *hrpf-1*. Lifespan curves of *ifg-1* mutants with an additional splicing factor knocked down. Approximately 30 worms were used in each trial. EV: empty vector RNAi bacteria; *ifg-1*: initiation factor 4G (eIF4G) family; *hrp-2*: heterogeneous nuclear ribonucleoprotein (HnRNP) R homolog; *hrpf-1*: HnRNP F homolog.

Table D.1: Gene function lifespan assay statistical data for RNAi originally thought to be targeting *hrp-2* and *hrpf-1*

Gene Name	Animals tested	Animals dead	Animals censored	Mean lifespan (Days)	% Mean lifespan	p value
Lifespan Trial 1						
N2/EV	48	34	14	21.34 ± 0.65	~	~
N2/*hrpf-1*	29	19	10	20.54 ± 0.68	96.25	0.3249
Lifespan Trial 2						
N2/*EV*	36	33	3	19.88 ± 1.02	~	
N2/*hrp-2*	31	28	3	19.89 ± 0.89	100.05	0.5678
N2/*hrpf-1*	30	24	6	22.10 ± 0.75	111.17	0.4467
ifg-1/EV	44	38	6	21.92 ± 1.18	110.26	
ifg-1/hrp-2	23	21	2	17.05 ± 1.05	77.78	0.0488
ifg-1/hrpf-1	37	33	4	20.62 ± 1.24	94.07	0.4375
Lifespan Trial 3						
N2/EV	118	104	14	18.77 ± 0.39	~	~
N2/*hrp-2*	78	70	8	18.95 ± 0.49	100.96	0.6825
N2/*hrpf-1*	22	17	5	18.71 ± 1.38	99.68	0.4111
ifg-1/EV	125	114	11	19.62 ± 0.61	104.53	~
ifg-1/hrp-2	153	141	12	20.34 ± 0.62	103.67	0.0418
ifg-1/hrpf-1	65	58	7	19.84 ± 0.96	101.12	0.2968
Lifespan Trial 4						
N2/EV	55	47	8	20.24 ± 0.54	~	~
N2/*hrp-2*	52	45	7	21.77 ± 0.65	107.56	0.0403
N2/*hrpf-1*	49	44	5	21.81 ± 0.67	107.76	0.0299
ifg-1/EV	31	26	5	19.35 ± 1.92	95.60	~
ifg-1/hrp-2	11	8	3	22.23 ± 2.63	114.88	0.0179
ifg-1/hrpf-1	19	19	0	21.11 ± 2.03	109.10	0.0454