Université de Montréal

Étude de la toxicité causée par le gène *C9orf72* dans la Sclérose Latérale Amyotrophique

Par Martine Therrien

Département de pathologie et biologie cellulaire

Faculté de médecine

Thèse présentée à la Faculté des études supérieures

En vue de l'obtention du grade de Ph.D.

En pathologie et biologie cellulaire -option système nerveux

Janvier 2016

RÉSUMÉ

La Sclérose Latérale Amyotrophique (SLA) est une maladie neurodégénérative qui affecte les neurones moteurs. 10% des cas sont des cas familiaux et l'étude de ces familles a mené à la découverte de plusieurs gènes pouvant causer la SLA, incluant *SOD1*, *TARDBP* et *FUS*. L'expansion de la répétition GGGGCC dans le gène *C9orf72* est, à ce jour, la cause la plus connue de SLA. L'impact de cette expansion est encore méconnu et il reste à déterminer si la toxicité est causée par un gain de fonction, une perte de fonction ou les deux.

Plusieurs gènes impliqués dans la SLA sont conservés entre le nématode *Caenorhabditis elegans* et l'humain. *C. elegans* est un vers transparent fréquemment utilisé pour des études anatomiques, comportementales et génétiques. Il possède une lignée cellulaire invariable qui inclue 302 neurones. Aussi, les mécanismes de réponse au stress ainsi que les mécanismes de vieillissement sont très bien conservés entre ce nématode et l'humain. Donc, notre groupe, et plusieurs autres, ont utilisé *C. elegans* pour étudier plusieurs aspects de la SLA.

Pour mieux comprendre la toxicité causée par l'expansion GGGGCC de *C9orf72*, nous avons développé deux modèles de vers pour étudier l'impact d'une perte de fonction ainsi que d'un gain de toxicité de l'ARN. Pour voir les conséquences d'une perte de fonction, nous avons étudié l'orthologue de *C9orf72* dans *C. elegans*, *alfa-1* (*ALS/FTD associated gene homolog*). Les vers mutants *alfa-1(ok3062)* développent des problèmes

moteurs causant une paralysie et une dégénérescence spécifique des neurones moteurs GABAergiques. De plus, les mutants sont sensibles au stress osmotique qui provoque une dégénérescence. D'autre part, l'expression de la séquence d'ARN contenant une répétition pathogénique GGGGCC cause aussi des problèmes moteurs et de la dégénérescence affectant les neurones moteurs. Nos résultats suggèrent donc qu'un gain de toxicité de l'ARN ainsi qu'une perte de fonction de *C9orf72* sont donc toxiques pour les neurones.

Puisque le mouvement du vers peut être rapidement évalué en cultivant les vers dans un milieu liquide, nous avons développé un criblage de molécules pouvant affecter le mouvement des vers mutants *alfa-1* en culture liquide. Plus de 4 000 composés ont été évalués et 80 ameliore la mobilité des vers *alfa-1*. Onze molécules ont aussi été testées dans les vers exprimant l'expansion GGGGCC et huit diminuent aussi le phénotype moteur de ces vers.

Finalement, des huit molécules qui diminent la toxicité causée par la perte de fonction de *C9orf72* et la toxicité de l'ARN, deux restaurent aussi l'expression anormale de plusieurs transcrits d'ARN observée dans des cellules dérivées de patient *C9orf72*. Avec ce projet, nous voulons identifier des molécules pouvant affecter tous les modes de toxicité de *C9orf72* et possiblement ouvrir de nouvelles avenues thérapeutiques.

Mots clés: sclérose latérale amyotrophique; *C. elegans*; *C9orf72*; criblage de molecules; dégénérescence

ABSTRACT

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disorder affecting the motor neurons. 10% of the cases are familial and using those families, many genes were shown to be involved in ALS pathogenesis, including *SOD1*, *TARDBP* and *FUS.* The GGGGCC repeat found in the first intron of *C9orf72* is, to this day, the most common genetic cause of ALS. Many hypotheses have been speculated to explain the toxicity of the pathogenic GGGGCC repeat, including loss and gain of function mechanisms.

Many proteins involved in amyotrophic lateral sclerosis (ALS) are evolutionarily conserved in the worm *Caenorhabditis elegans*. *C. elegans* is a transparent nematode widely used for anatomical, behavioural and genetic studies. It possesses an invariant cell lineage that includes 302 neurons in the adult nematode. Also, cellular stress responses and survival mechanisms are genetically regulated and conserved from the nematode and human. Therefore, our group, and others, have used *C. elegans* to model different aspects of neurodegenerative diseases including ALS.

To better understand the toxicity caused by the GGGGCC repeat expansion in *C9orf72*, we have developed two *C. elegans* models to understand either the impact of the loss of function of *C9orf72* or the gain of toxicity of the RNA containing the GGGGCC repeat. To understand the loss of function, we have characterized the orthologue of *C9orf72* in *C. elegans*, *alfa-1* (*ALS/FTD associated gene homolog*). Mutant *alfa-1* worms exhibit motor impairments leading to paralysis and neurodegenereation of the GABAergic neurons. Mutant worms are also sensitive to osmotic stress which can lead to increased neurodegeneration. On the other part, exposure of *C. elegans* neurons to the

RNA containing the GGGGCC repeat causes also motor problem and degeneration affecting the motor neurons. Therefore, our data suggest that both loss of function of *C9orf72* and toxic gain of function are detrimental to neurons.

Since motor dysfunctions in worms can be easily accessed in liquid culture, we have screened more than 4,000 FDA approved compounds in the *alfa-1(ok3062)* worms. 80 molecules were shown to improve *alfa-1* impaired function and eleven of those were also tested for their effect to reduce the neurotoxicity caused by the GGGGCC repeat RNA. Eight molecules were shown to affect both types of neurotoxicity.

Finally, from these eight molecules that can improveboth types of toxicity, two were shown to restore the abnormal RNA expression observed in *C9orf72* patient-derive cells. With this project, we aimed to identify molecules that can affect the loss of *C9orf72* toxicity and the toxic gain of RNA function containing the GGGGCC repeat to hopefully open new therapeutic avenues for ALS patients.

Key words: Amyotrophic lateral sclerosis; *C. elegans*; *C9orf72*; drug screening; neurodegeneration

Table of content

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

- a.a.: amino acid
- ALS: amyotrophic lateral sclerosis
- BAC: bacterial artificial chromosome
- CRISPR/Cas9 clustered regulatory-interspaced short palindromic repeats/CRISPR associated protein 9
- DENN: differentially expressed in normal and neoplastic cells
- DM1: myotonic dystrophy type 1
- DNA: deoxyribonucleic acid
- EMS: ethyl methanesulfonate
- ER: endoplasmic reticulum
- fALS: familial amyotrophic lateral sclerosis
- FDA: food and drug administration
- FTD: frontotemporal dementia
- GFP: green fluorescent protein
- HSP: hereditary spastic paraplegia
- IIS: insulin-IGF signaling pathway
- iPSC: induced pluripotent stem cell
- IRES: internal ribosomal entry site
- mRNA: messenger ribonucleic acid
- PCR: polymerase chain reaction
- RAN translation: repeat associated non-ATG translation
- RNA: ribonucleic acid
- RNAi: interference ribonucleic acid

- ◊ sALS: sporadic amyotrophic lateral sclerosis
- ◊ SCA2: spinocerebellar ataxia type 2
- ◊ SMA: spinal muscular atrophy
- ◊ TMP: 4,5',8-trimethylpsoralen
- ◊ UTR: untranslated region

It matters that you don't just give up

Stephen Hawking

ACKNOWLEDGMENTS

Wow it's almost over already!

I heard somewhere that we are the average of the 5 people we hang out the most with. I don't know where I stand among those 5 people but I can say that I was very lucky to be surrounded by very talented people during this journey. I would like to thank all members of the Rouleau's and Parker's laboratories. Thank you for teaching me, for sharing with me your knowledge and for helping me. Special thank to Guy Rouleau who was not afraid to welcome a CEGEP student in his lab 10 years ago. Thanks for giving me this great opportunity. Also, I would like to thank Patrick and Alex. You helped me built this project and I'm very proud of what I have accomplished with you. Thanks for supporting me and trusting me. I think you both share this enormous curiosity toward science and it is always great to work with people that love what they do. I have learnt so much from Alex's calmness and Patrick's crazy ideas…because they are not that crazy most of the time. I'll miss talking science with both of you.

I have spent 10 years in the Rouleau lab and so many things have happened during those years. The sad part of our field is that if we do well, people come and go very fast. My road crossed the ones of many people and I am grateful to have met you all. I have shared much laughter, memories, drinks and sweet treats with many of you. I leave this project in the good hands of James and Amélie, good luck to both of you, I know you will succeed. Throughout the years, some colleagues became friends, very good friends. Hélène, Pascale, Loubna, Sandra and Claire, thanks for always having time for me, support is important is this big lab. Thanks Kathrin, life needs this little sparkle that you have. Very special thank to Simon and Anaïs. Thanks for understanding how weird I can

be…probably because you also are. A Ph.D is not always easy and you were there in the good and in the less good moments, thanks!

On a more personal side, I would like to thank my 'non-scientist' friends; Marlène, Laurence, Carol, and my family; Nicolas, Carole and Guy. Thanks for supporting me even though you did not always understand why I really needed to count worms at odd hours. I truly believe that life is about balance and as much as I like what I do, I am grateful that you reminded me sometimes that it was time to take a little break from the lab. Without knowing it, you were part of this journey and you made me better at it.

Merci énormément !

Martine

CHAPTER 1: INTRODUCTION[1]

ALS

DISEASE INCIDENCE AND DEVELOPMENT

Amyotrophic lateral sclerosis (ALS) is a fatal, late onset, neurodegenerative disorder. It is speculated that most cases of ALS are sporadic (sALS) (~90%), but ~10% are familial (fALS). Many groups have suggested that these numbers are probably an underestimation of the fALS cases [3,4]. There is no definitive criterion for fALS but the general consensus is that the presence of ALS in either a first or second degree relative constitutes the familial form of the disease [5]. ALS cases with no known family history are therefore referred to as sALS. However, sALS and fALS are clinically indistinguishable [4].

The disease is characterized by the loss of motor neurons in the brainstem and spinal cord leading to muscle weakness, fasciculation and wasting [6]. Symptoms typically start in a specific region of the body, causing either bulbar (25%), cervical or lumbar onset (75%), which specifically characterizes the type of ALS. Ultimately most motor neurons become affected and death by respiratory failure usually occurs 3-5 years after the onset of symptoms when the respiratory muscles are denervated.

---

[1] This chapter is inspired by literature reviews written by Martine Therrien, including [1], [2] and Therrien et al. Current Neurology and Neuroscience report, submitted

Incidence of ALS in western countries is 2-16/100 000 individuals but it affects people worldwide [6]. Men are more at risk than women to develop sALS, but the risk ratio is the same for fALS cases [6]. The age of onset varies between 47-63 years old with a slightly earlier age of onset in fALS cases; however the incidence decreases drastically after the age of 80 years [6].

Diagnosis of ALS is usually a long process and a definite diagnosis is often only made when the patient's symptoms do not fit the symptoms of any other possible conditions. ALS can be mistaken with other motor neuron or nerve diseases or lesions; neuromuscular junction disorders; or myopathies[6]. Clinicians base their diagnostics on the presence of upper and lower motor neuron signs in the same region of the body[6] and the lack of neuroimaging, electrophysiological and pathological evidences of other diseases that could explain the symptoms [7]. According to the '*El Escorial*' and 'Airlie House' diagnostic criteria, which include the main elements for the classification of ALS patients, to obtain a 'definite' ALS diagnosis, patients must have clinical evidences of upper and lower motor neuron signs in three distinct regions of the body. At diagnosis, only 40% of patients with a family history of ALS fit the 'definite' diagnostic criteria while at death, 10% of all patients still remain with only a 'possible' diagnostic of ALS [8].

TREATMENT

Riluzole, an inhibitor of glutamate release, is the only drug presently used to alleviate the symptoms of ALS patients [6]. It has a modest effect, increasing survival from 3-6 months. Therefore, symptomatic treatment is the sole alternative for patients. Multidisciplinary

teams of physiotherapists, occupational therapists, respiratory physicians and neurologists guide patients throughout the development of the symptoms and this type of therapy was shown to reduce the risk of death by 45%, within 5 years [6]. Finally, weight loss and respiratory failure are the main outcomes in ALS patients, thus the use of a ventilator and percutaneous gastrostomy are used to manage symptoms and delay death.

GENETICS

Most cases of ALS are sALS (90%), but 10% are fALS and exhibit a Mendelian pattern of inheritance [9]. The study of fALS has led to the identification of many disease causative genes [10] that explain fALS and sALS cases (Figure 1.1 and Table 1.1).

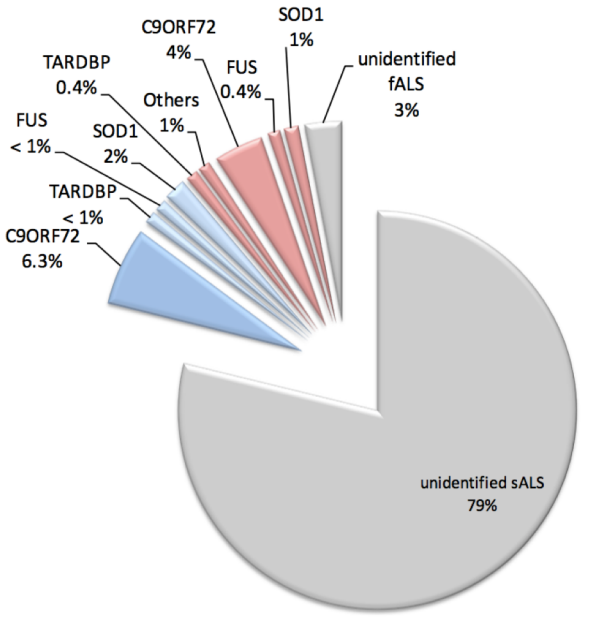


Figure 1.1: Genetic causes of ALS. In red are the fALS cases and blue the sALS explained by a genetic mutation. In grey are the sALS and fALS cases with unknown cause

In 1993, mutations in the gene *SOD1 (Superoxide dismutase 1)* were linked to many fALS cases [11] and a few unrelated cases of sALS [12]. Since then, more than 160 mutations have been linked to sALS and fALS [13]. SOD1 is a ubiquitous enzyme that catalyzes the removal of superoxide into oxygen and peroxide. Mutations are thought to affect the folding of the protein causing a toxic gain of function of the mutant proteins that accumulate at the mitochondria [14,15]. Even though it is not the only possible mechanism, it is speculated that mitochondrial dysfunction is central to *SOD1* pathogenesis [15].

The identification of TDP-43 as the main constituent of the aggregates found in post-mortem patient spinal cord and brain tissues [16] led to the identification of mutations in *TARDBP (Tar DNA binding protein)*, the gene encoding TDP-43 protein [17,18]. Since then, more than 40 mutations in this gene were shown to be causative of fALS and sALS [13,17,18]. TDP-43 is a ubiquitous protein that contains two RNA binding domains, a glycine rich domain and nuclear import and export signals [19]. In the wild-type state, TDP-43 shuttles between the cytoplasm and the nucleus, but its specific function is unknown. It was identified as an important player in various aspects of RNA processing, including, transcription, translation, splicing and micro RNA processing; and involved in RNA transport and stress granules formation [19]. Shortly after the identification of *TARDBP,* mutations in *FUS* (*Fused in sarcoma*) were found in ALS patients [20,21]. FUS is also a RNA binding protein that shares many functions with TDP-43 [22]. It is speculated that mutations in *FUS* and *TARDBP* cause a toxic gain of function of the mutant proteins that relocalize in the cytoplasm, but much remains to be understood about their specific role and toxicity in motor neurons.

With the fast evolution of sequencing techniques, many genes were recently linked to ALS [10] (Table 1.1). To this day, over twenty genes have been speculated to be involved in ALS pathogenesis. Studies to clearly evaluate the importance of the individual genes have been undertaken [23,24] but large studies are required to unambiguously confirm their contributions to ALS pathogenesis in different subtypes of patients and in different regions of the world.

Human gene	Function	*C. elegans* orthologue	*C. elegans* transgenic model
Most common ALS genes			
C9orf72	Unknown, DENN protein	*alfa-1*	*Na*
FUS	RNA metabolism	*fust-1*	*Yes*
OPTN	Autophagy	---	*Na*
PFN1	Cytoskeleton dynamic	*pfn-1*	*Na*
SOD1	Superoxide metabolism	*sod-1*	*Yes*
SQSTM1	Autophagy	*sqst-2*	*Na*
TARDBP	RNA metabolism	*tdp-1*	*Yes*
UBQLN2	Autophagy	*ubqnl-1*	*Na*
VCP	Autophagy	*cdc-48.1/2*	*Na*
Other genes involved in ALS			
ALS2	Endocytosis	---	*Na*
ANG	Blood vessels formation	---	*Na*
ARHGEF28	RNA metabolism	*rhgf-1*	*Na*
CHCHD10	Mitochondrial function	*har-1*	*Na*
CHMP2B	Autophagy	---	*Na*
DCTN1	Cytoskeleton dynamic	*dnc-1*	*Na*
FIG4	Vesicular trafficking	*C34B7.2*	*Na*
GLE1	RNA metabolism	---	*Na*
HNRNPA1	RNA processing	*hrp-1*	*Na*
HNRNPA2B1	RNA metabolism	*hrp-1*	*Na*
MATR3	RNA processing	---	*Na*
SETX	RNA processing	---	*Na*
SPG11	DNA damage	---	*Na*
TBK1	Protein degradation	---	*Na*
VAPB	Vesicular trafficking	*vpr-1*	*Na*

Table 1.1 ALS genes. Genes linked to ALS, their *C. elegans* orthologues and a summary of the transgenic *C. elegans* models published. The importance of the gene is based on [23,24]

Sequencing and resequencing of large ALS cohorts have led to the identification of many individuals with more than one disease causative mutation (reviewed by Lattante *et al* [25]). These mutated genes might reflect low penetrance genes, or genes that cannot induce ALS alone. However, it cannot be excluded that those genes interact to vary disease onset, progression and symptoms.

New genetic concepts are emerging to explain some sALS cases. One of the main hypotheses explaining sALS cases was the impact of the environment in the development of the disease [26] . Groups have studied the impact of smoking, exercise and exposure to heavy metals and pesticides in different ALS cohorts (reviewed by Al-Chalabi & Hardiman [26]). However, epidemiological studies like these can be challenging and few are reliable, most of the time because of the lack resources available to do the study [26]. Therefore, no specific environmental factors have been found thus far to clearly explain sALS cases.

With the rapid advancement of sequencing technologies, new genetic concepts have emerged to explain sALS. Amongst these is the concept of *de novo* mutations which was previously shown to be a key player for many neurodevelopmental and psychiatric diseases [27,28]. *De novo* mutations arise during the fertilization of germline and results in mutations found in the offspring that were absent in the parents, therefore causing sALS with a genetic cause. Mutations in a few known ALS genes were found in some sALS patients [29,30]. *FUS* is the gene that is the most affected by *de novo* mutations so far. Many novel and previously identified mutations affecting the coding region and splicing of *FUS* were linked to early onset sALS [31-34], hence suggesting that *de novo* mutations can have a role in sALS cases.

Many pathological pathways were speculated to play a role in the pathogenesis of ALS. At the functional level, many of the genes linked to ALS share common cellular roles. RNA processing, mitochondrial dysfunction, stress response and protein degradation have emerged to be important pathological concepts. However, for most of the genes, the exact cellular function and the impact of the mutant proteins still remain to be established. It is still necessary to elucidate how these mutant genes can be specifically detrimental to motor neurons and how disease progression can vary among individuals with the same mutation.

PATHOLOGICAL CHARACTERISTICS OF ALS

Aside from the motor neurons, which cells contribute to ALS pathogenesis is still under evaluation. TDP-43 inclusions are observed in glia and neurons of the motor cortex, but also in the brainstem, the spinal cord and in white matter [35]. Based on the pathological observation that protein inclusions are also present in glial cells, different groups have speculated that these non-neuronal cells may participate in the death of the motor neurons. Using *SOD1* mouse models, it was shown that onset of symptoms require expression of mutant protein in the motor neurons but the mutant protein also had to be expressed in non-neuronal cells such as astrocytes and microglia to affect to disease progression [36,37]. Therefore indicating that dysfunction of different cell types might be involved in ALS pathogenesis. However, thus far these types of *in vivo* observations were only made for *SOD1*. A few *in vitro* studies have been done with mutant TDP-43 toxicity and suggested

opposite results [38,39] therefore, more studies are required to examine the contribution of non-neuronal cells in ALS pathogenesis.

Protein aggregates are often observed in neurodegenerative disorders and it is also the case for ALS. TDP-43 is the main constituent of the aggregates observed in motor neurons of fALS and sALS [16]. Protein aggregates containing FUS, p62, SOD1, UBQLN2 and *C9orf72* dipeptides repeat proteins are also observed in post-mortem, brain and spinal cord tissues of different subset of ALS patients [7,40]. Protein aggregates are observed in motor neurons of the motor cortex, brainstem and spinal cord as well as in glial cells in those regions [35]. It is interesting to note that some aggregates are mutually exclusive suggesting independent toxic pathways in motor neurons[7]. However, it is still unclear if these aggregates are toxic or protective. Many proteins that were found in aggregates as well as the proteins encoded by some ALS genes were shown to participate to protein homoeostasis and could therefore affect the formation and/or clearance of these aggregates [41,42]. Abnormal protein homoeostasis leading to the presence of aggregates or inclusions is a hallmark of many neurodegenerative disorders so it is difficult to know if these are specific to ALS or are a general feature of age-related neuronal death.

SIMILARITIES BETWEEN ALS AND OTHER NEURODEGENERATIVE DISORDERS

FRONTOTEMPORAL DEMENTIA (FTD)

FTD is a group of non-Alzheimer dementia characterized by atrophy of frontal and/or temporal lobes leading to behavioural changes or language decline [43]. It is characterized by

pathological protein inclusions in which either TDP-43, TAU or FUS proteins are found in the affected regions of the brain [43]. Most cases are sporadic but 10-20% of cases have a genetic component. Mutations in *MAPT (microtubule-associated protein tau)*, *GRN* (*Granulin*) or *C9orf72* are linked to FTD and are the most common genetic causes of FTD [43].

For many years, it was speculated that ALS solely causes motor dysfunction, however, it is now established that subsets of ALS patients also exhibit cognitive deficits. Patients' cognitive dysfunctions range from mild cognitive deficits to dementia fitting the criteria of FTD. 50% of ALS patients have been shown to develop cognitive deficits [44,45] and 15-20% were shown to fit the criteria for FTD. Also, up to 27% of FTD patients exhibit motor dysfunction [46]. Therefore, it is now accepted that ALS and FTD are along the same pathogenic continuum where patients range from pure motor neuron symptoms to pure dementia with many individuals found to have a mixture of both. Mutations in *C9orf72*, *TARDBP*, and *UBQLN2* were found to be causative of ALS and FTD [17,18,47-49]. It is still unclear how in a single family, the same genetic variant can cause either ALS and FTD (for an example of *C9orf72* [50]). Therefore, much remains to be investigated about the similarities and differences leading to ALS or FTD pathogenesis.

Polyglutamine Disorders

ALS proteins are also found in aggregates of other neurodegenerative disorders [51-53]. An example is the presence of TDP-43 and FUS proteins in aggregates of polyglutamine disorders. Those proteins have been shown to co-localize to the polyglutamine aggregates in Huntington's Disease in cells and tissue mouse model, and in post-mortem tissues of

patients affected by Machado-Joseph Disease [54,55]. FUS and TDP-43 were also shown to interact with ATXN-2 (encoded by *ATXN2(Ataxin-2)* gene) [56,57] a protein that contains a polyglutamine tract that has been shown to cause Spinocerebellar ataxia type 2 (SCA2) when the polyglutamine repeat is longer than 34 units [58]. Intermediate CAG repeat (*i.e.* repeat length around 28 units) found in *ATXN2*, have been shown to be a modifier of pathogenesis in TDP-43 model organisms and sALS patients [56] , highlighting the special relationship between ALS and CAG repeat disorders.

COMMON PATHWAYS IN NEURODEGENERATIVE DISEASES

Neurodegenerative disorders share many common pathological pathways. Impaired protein degradation, endoplasmic reticulum (ER) stress and neuroinflammation are some examples.

Many neurodegenerative diseases are characterized by the composition of the protein aggregates that are observed in tissues of affected individuals. Alzheimer's disease is characterized by the presence of amyloid plaques and tau filaments [59], Parkinson's disease is characterized by the presence of Lewy bodies, which are composed mainly of alpha-synuclein protein [60] and most polyglutamine diseases are characterized by inclusions containing the expanded polyglutamine tract[61]. As mentioned above (section *Pathological characteristics of ALS*), impaired protein homoeostasis is a key feature of ALS pathogenesis. Therefore, abnormal protein synthesis or degradation leading to the formation of cytoplasmic or nuclear protein aggregates is a recurrent theme in neurodegeneration. However, the role that these aggregates play in pathogenesis is still unclear.

Also, ER stress has been a recurrent theme regarding ALS pathogenesis in many models as well as in patients [62] . In addition, neuroinflammation, was observed in patient post-mortem tissues and has been shown to alleviate phenotypes observed in different model organisms [63]. Similar toxic mechanisms were also implicated in Parkinson's disease and Alzheimer's disease [64-67]. These data suggest that neurodegenerative disorders share many common toxic pathways and insight into a specific disease could be gained by comparing them.

C9ORF72

GENETICS

Many linkage studies using families affected by ALS and/or FTD from different areas of the world identified a region on chromosome 9p as a common genetic cause of ALS [68-70]. In 2011, two independent groups simultaneously identified a causative ALS gene on chromosome 9p as being the non-coding repeat expansion GGGGCC found in the first intron of the *C9orf72* gene. *C9orf72* generates three alternatively spliced transcripts; two of which (V1 and V3) produce the long protein form of *C9orf72* (481 a.a) and the other transcript (V2) produces a shorter form (222 a.a.) (Figure 1.2). The GGGGCC repeat is located in the promoter region of the V1 isoform and the first intron of the two others (Figure 1.2).

In the first reports, the size of the repeat was suggested to be less than 10 units in

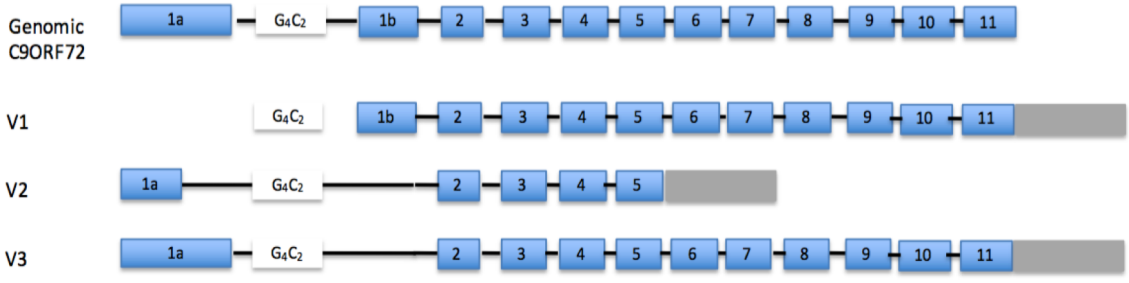


Figure 1.2: *C9orf72* gene and isoform structure. V1 refers to transcript NM_145005, V2 refers to transcript NM_018352 and V3 refers to transcript NM_1256054

healthy individuals. Its size in affected individuals appeared to be at least 30 units; while it was already clear that the number of repeats could reach hundreds and even thousands of units in patients [47,48]. More recent reports have since then identified some ALS and/or FTD cases with repeat size as small as 22 units [71,72]. Nonetheless several studies have now

reported healthy individuals with 20-30 repeats [71-73] so the minimal repeat length that can trigger ALS/FTD still remains to be precisely established.

Following these first reports, many others have followed and identified the GGGGCC expansion as being a major cause of ALS and/or FTD in individuals from European countries (Finland[74], France[75], United Kingdom, Ireland[76], Spain[77], Italy[78], Holland[79], Grece[80], Portugal[81], Belgium and Poland), the Americas (USA[82], Canada[50], Brezil[83]) and Australia [84]. Most individuals are heterozygous for the expansion, carrying an expanded and a wild-type allele, but a few cases were identified to be homozygous [85,86]. In one homozygous case, disease severity was stronger than what is typical of FTD cases [86] but in the other one, the disease progression and severity of ALS/FTD symptoms were similar to typical ALS/FTD cases [85]. Interestingly, with the exception of the Kii peninsula of Japan [87], most studies done in Asian populations have shown that *C9orf72* is not a major genetic cause of ALS in this area of the world (Korea [88,89], China [90,91], Japan [92-94], Taiwan [95]).

Even though the ALS/FTD symptoms have a variable age of onset, several groups reported almost complete penetrance of the expansion by the age of 80 [96,97]. However, a few individuals were reported with an expanded GGGGCC repeat and did not develop neurodegenerative symptoms after this age [92,98] .

CHARACTERISATION OF REPEAT LENGHT

Since the expansion in *C9orf72* is a pure G/C repeat, its quantification has been challenging. Repeat primed PCR can accurately only detect repeats smaller than 60 units and Southern

blot analysis requires large amounts of DNA so it has not been carried out for most individuals. Also, when using the same DNA samples, groups using repeat primed PCR to quantify the repeat length often misquantified it, suggesting that interpretation of repeat size should be done carefully, if not exclusively done by Southern blotting [99].

Anticipation is a concept by which the number of repeats can expand in size from one generation to the next and correlates with an increase in disease severity and an earlier age of onset. This phenomenon is observed in most coding and non-coding repeat disorders [100]. Some researchers have speculated a link between age of onset or survival and repeat size in *C9orf72*[96,101]. Also, one group reported major anticipation in *C9orf72* patients [102], with disease affecting children 7 years younger than the age at which their parents became affected. Nonetheless, other independent teams could not replicate these findings and the importance of anticipation in *C9orf72* toxicity remains a matter of debate [47,48,74,103,104] . The small sample sizes might be the reasons for these conflicting observations.

Most studies that evaluated the repeat length have used genomic DNA prepared from peripheral blood. However, in repeat disorders it is often observed that the repeat length varies across different tissues [105]. Many reports showed similar results for the GGGGCC expansion of *C9orf72* and neuronal cells seem to exhibit an increase in repeat length when compared to non-neuronal cells of the same individual [104,106,107] . Therefore suggesting that the repeat length expressed in neurons is probably underestimated.

OTHER MUTATIONS

Aside from the GGGGCC repeat found in the first intron of *C9orf72*, different groups have tried to identify mutations outside this region in ALS and FTD patients. Different groups have confirmed that most expansion carriers share a common haplotype that includes 10 nucleotides surrounding the repeat [47,92,94,108].

Outside the expansion region, a 10 bp deletion has also been identified in the first intron in a few FTD patients [109] and some missense variations were found in the coding region of *C9orf72* in sALS [110]. However, the pathogenic contribution of these variants is still unclear.

C9ORF72 EXPRESSION

EXPRESSION OF *C9ORF72* IN MICE

The expression of *C9orf72* was studied in mice by using a Lac-Z insertion in the mouse *C9orf72* sequence [111] and by *in situ* hybridization [112]. In adult animals, expression was found in the brain, spinal cord, spleen, kidney and testes [112], whereas no expression was observed in muscle, heart, lungs and liver [111]. Expression of mouse *C9orf72* during development was reported in neurons and different organs during embryonic stages [112]. mRNA and protein expression of the different isoforms was shown to change during development [113] . For example, isoform 1 is the most expressed at postnatal day 1 and is found mainly in the nucleus. Isoform 2, however, is expressed mainly in the cytoplasm at postnatal day 1, but its expression increases in the nucleus, during development, to reach

its maximum at postnatal day 56 [113]. In primary mouse culture, cortical neurons show expression of *C9orf72* in neurites and growth cones [113] . Thus, experiments in mice show that *C9orf72* is expressed during development and adulthood in the nervous system and suggest that the different isoforms could have different roles during development.

EXPRESSION OF *C9ORF72* IN HUMANS

In humans, expression of *C9orf72* was studied with *in situ* hybridization probes and showed expression in neuronal cells of the spinal cord [111]. Using qRT-PCR, expression of the *C9orf72* isoforms 1 and 2 was shown in cervical spinal cord, cerebellum and motor cortex in post-mortem tissues [114]. In the central nervous system, the highest expression of *C9orf72* occurs in the cerebellum and the lowest in the putamen [48]. Expression in cultured fibroblasts is low and seems to increase when the cells are used for the differentiation of iPSC-neurons [114,115], suggesting an important tissue specific expression of *C9orf72*.

Few antibodies are commercially available to clearly detect C9ORF72 protein, however, one study has confirmed C9ORF72 protein expression in frontal and cerebellar cortex [116]. Novel antibodies were recently generated to detect the long and short C9ORF72 isoforms and confirmed the expression of both isoforms in motor, temporal and frontal cortexes as well as cerebellum and lumbar spinal cord [117].

The exact cellular expression and localization of *C9orf72* in humans still need to be elucidated, but preliminary results suggest localization at the nuclear membrane for the short isoform, and in mischaracterized cytoplasmic puncta for the long isoform [117]. It is

interesting to mention that in Western blotting, the isoforms were found in different solubility fractions, where the long isoform seems to be more insoluble than the short one [117], suggesting that the cytoplasmic puncta might be insoluble and less dynamic.

In conclusion, expression analysis of *C9orf72* in mice and humans has confirmed the importance of this protein in the nervous system during development and adulthood. Furthermore, even though these results remain to be confirmed, expression analysis suggests a role for C9ORF72 protein in the cytoplasm and at the nuclear membrane.

FUNCTION OF C9ORF72

Very little was known about the regular function of *C9orf72* when it was first linked to ALS. Bioinformatic analysis showed that the full sequence of C9ORF72 protein shares many characteristics with DENN proteins (Differentially Expressed in Normal and Neoplastic cells) [118,119]. DENN proteins are highly conserved guanine nucleotide exchange factor (GEF) proteins involved in endocytosis and intracellular trafficking [120]. In motor neurons, C9ORF72 protein was shown to colocalize with some RAB proteins and regulates autophagy, partially confirming that *C9ORF72* can act as a DENN protein [121].

Recently, different *in vivo* models suggested that the GGGGCC expansion can affect RNA export from the nucleus and different modifier screens conducted in model organisms have shown that nuclear import and export proteins could modify this toxicity [122-124]. These data are also supported by the fact that the short protein isoform of C9ORF72 was shown to be located at the nuclear membrane [117], suggesting a function at *C9ORF72* at the nuclear

membrane. However, a recent study also suggested that the presence of cytoplasmic aggregates alone could affect nucleo-cytoplasmic shuttling independently of the function of the proteins found in the aggregates [125]. Therefore, knowing that *C9orf72* positive patients exhibit many protein aggregates, the toxicity caused by the expanded GGGGCC and the function of C9ORF72 protein at the nuclear membrane still remains to be confirmed.

Additionally, the pathogenic GGGGCC repeat expansion was shown to affect the formation of stress granules [121,126,127]. Even though other ALS proteins, such as FUS and TDP-43, were also shown to participate in stress granule formation (reviewed by Ling *et al* [128]), little is known about this function of *C9orf72* in regards to ALS pathogenicity.

The functions of *C9orf72* in normal and disease states remain to be elucidated in different cell types. However, it seems to affect RNA metabolism and endosomal trafficking, two important pathways that were previously shown to be involved in ALS pathogenesis.

CONSERVATION OF C9ORF72 ACROSS SPECIES

Model organisms are important tools to learn about the function of new proteins. Therefore, conservation of *C9orf72* in different species was examined. The GGGGCC repeat found in the first intron of the gene has only been found in primates and it is not conserved in mice or lower model organisms. However, the gene is highly conserved (over 90% identity) in chimpanzees and marmosets [111]. In lower organisms, most amino acids are conserved in mouse, rat, chick embryo, zebrafish with between 66%-98% identity in these

organisms [111]. Most conserved residues are distributed across the protein suggesting that its function is conserved across the species [118].

REPEAT DISORDERS

ALS is not the only disease that can be caused by a nucleotide repeat expansions. Nucleotide repeats represent 30% of the human genome and vary in length and frequency [129]. Many repeats were shown to cause neuro-developmental or neurodegenerative disorders when the repeat exceeds a certain threshold [130]. Repeats can be found in the coding or non-coding regions, introns or UTRs, of the affected genes and cause either a toxic loss and/or gain of function mechanism. Here are a few examples of well-studied repeat disorders and how those repeats can induce cellular toxicity.

MYOTONIC DYSTROPHY AND THE TOXICITY OF RNA FOCI

Myotonic dystrophy type 1 (DM1) is a common repeat disorder and the most common muscular dystrophy. It is clinically similar to myotonic dystrophy type 2. The onset of symptoms can be observed during birth, childhood or at adulthood leading to a wide variety of symptoms including mental retardation, muscle degeneration, heart defects and cataracts [131]. It is caused by a CTG repeat in the 3' UTR of *DMPK1* (*Dystrophia myotonica protein kinase*) gene [132-134]. Healthy individuals have between 5-35 units of the CTG repeat, whereas patients have at least 50 units. Repeat can be as long as thousands of units and disease severity correlates with repeat length.

Most of the research concerning the toxicity of RNA foci arises from studying DM1. The transcript containing the expanded repeat causing DM1 accumulates in the nucleus of muscle cells. Interestingly, it was shown that those RNA foci sequester muscleblind-like (MBLN), a protein that binds to CUG repeats [135], leading to its loss of function. Knock-out

mice of *Mbln* recapitulates features of DM1 including myotonia, heart defects and abnormal splicing [135] suggesting that toxicity observed in DM1 reflects the loss of function of *Mbln* in muscle cells. Therefore, RNA foci can be toxic to cells by sequestrating proteins.

POLYGLUTAMINE DISEASES AND ABNORMAL TRANSLATION

One of the most common toxic nucleotide repeat is the CAG expansion. CAG codon encodes for glutamine and coding and non-coding CAG pathogenic expansions were found in a dozen genes causing a wide variety of symptoms [136]. While the genetic causes underlying these diseases are known, the toxic mechanisms involved are unclear and no treatment is available. Aside from the CAG repeat encoding gene, genes encoding a CTG repeat were found to be transcribed in the anti-sense direction leading to the production of toxic transcript containing CAG and CUG codons [137] therefore increasing the spectrum of disorders caused by CAG repeats. The presence of a long C/G rich repeat in a transcript can lead to the formation of an abnormal secondary structure [138]. Knowing that translation is highly sensitive to RNA secondary structure, researchers have studied the impact of expanded toxic repeat on translation.

The gene *ATXN3* (*Ataxin-3)* contains a coding CAG repeat in its 3' end . When the repeat reaches 55 units, it causes Machado-Joseph disease, also called Spinocerebellar ataxia type 3 (SCA3), a neurodegenerative disorder causing the loss of Purkinje cells in the cerebellum [139]. The CAG repeat causes the production of toxic long polyglutamine tracts. However, when the polyglutamine tract is encoded by CAA codons, which also leads to the production of polyglutamine, no toxicity is observed[140-142]. Therefore, it was hypothesized

that a change in reading frame, called ribosomal frameshifting, occurred along the CAG tract during translation leading to the production of polyalanine instead of polyglutamine. This phenomenon was shown in cell models and model organisms [141,142]. Ribosomal frameshifting is widely characterized in viruses that use this method to increase the efficiency of their genomes by encoding more than one protein from a single RNA transcript [143].

Another method used by viruses to increase the efficiency of their genome is called internal ribosomal entry site (IRES)[144]. It was shown that the secondary structure of the RNA that contains a CAG repeat can attract the cellular translation machinery and initiate translation in a non-ATG manner, a process called repeat-associated non-ATG translation (RAN translation). Even though RAN translation does not act exactly like an IRES, it was speculated that the structure formed by the repeat can act in a similar manner[145]. Interestingly, the phenomenon was shown to happen along the *ATXN8 (Ataxin-8)* and *HTT* (*Huntingtin)* transcripts, causing Spinocerebellar ataxia type 8 and Huntington's disease, respectively, when the CAG repeats that is encoded in the *ATXN8* and *huntingtin* genes are above a certain threshold [146,147]. When RAN translation occurs, expression of polyglutamine is independent of the presence of the ATG starting codon upstream of the repeat [148,149]. Since no ATG is used, translation can occur in all reading frames along the sense and anti-sense transcripts, leading to the production of numerous polypeptides from a single CAG repeat containing transcript. This phenomenon was also observed along the CGG non-coding repeat of *FMR1(Fragile X mental retardation 1)* gene, the genetic cause of Fragile-X-associated tremor ataxia, leading to the production of polyglycine and polyalanine peptides [150].

FRAGILE X SYNDROME AND TOXICITY OF A DECREASE OF GENE EXPRESSION

The CGC repeat found in 5' UTR of *FMR1* is an example of how a repeat expansion can affect its own gene expression. In healthy individuals, the repeat length varies between 6-55 units. When the expansion is longer than 200 units, its causes a neurodevelopmental disorder called Fragile-X syndrome [151]. The disease is characterised by mental retardation, as well as behavioural and social problems similar to autism spectrum disorders [136]. FMRP, the protein encoded by *FMR1*, is an RNA binding protein that shuttles between the nucleus and the cytoplasm and directly binds to mRNA. In Fragile-X syndrome patients, a hypermethylation of the CpG islands located in the 5' UTR has been observed causing a loss of *FMR1* expression [151]. Interestingly, a knock-out mouse model of *FMR1* exhibits many features similar to patient symptoms including hyperactivity, anxiety behaviours and cognitive deficits [136]. Therefore, these data suggest that a decreased expression of *FMR1* is sufficient to induce the phenotypes observed in Fragile-X syndrome.

Interestingly, if the CGC repeat length is intermediate (*i.e.* between 70-200 units), expression of the *FMR1* RNA is elevated and was shown to cause a neurodegenerative disease called Fragile-X tremor/ataxia syndrome. This disease is characterised by loss of Purkinje cells and atrophy of the cerebellum. Intranuclear aggregates containing proteins and the *FMR1* mRNA are observed in patients post-mortem tissues and are speculated to cause toxicity by sequestrating other RNA binding proteins, but the toxicity of these RNA foci is not as clear as for DM1[152].

C. ELEGANS

CURRENT MODELS TO UNDERSTAND ALS

With more than 20 genes now linked to ALS, the consequences of the mutant proteins need to be evaluated. Since the identification of *SOD1* in 1993, many mouse models have been established to understand the function of the mutant proteins in normal and disease states. Transgenic mouse models expressing mutant *SOD1* were able to successfully model the ALS progression and the motor neuron phenotypes observed in patients (among others[153-155]).However, *SOD1* mutations are found in only a small fraction of ALS patients. Even though some models seem promising to understand different aspects of ALS pathogenesis, transgenic animals expressing either TDP-43 and FUS proteins cannot recapitulate the involvement of these proteins in motor neuron integrity as well as the *SOD1* models. By itself, expression of wild-type TDP-43 in mouse causes neuronal loss, decreased survival and pathological characteristics of ALS [156,157]. When expressing mutant TDP-43 proteins, some models exhibit phenotypes related to ALS pathology, motor phenotypes, decreased survival and astrogliosis, but surprisingly, many models exhibited only minimal neuronal loss and TDP-43 aggregates [158-163] . Expression of wild-type and mutant FUS in mouse spinal cord can recapitulate pathological characteristics of ALS but, again, in neither cases was neuronal loss observed [162,164]. Also, ALS mouse models have failed over the past years to correctly predict the efficacy of drugs that were tested in clinical trials, which led many researchers to speculate that alternative approaches should be developed to evaluate potential drugs [7,165-167]

In recent years, some groups have started using human derived induced pluripotent stem cell (iPSC) derived neurons to study ALS pathogenesis. Since these iPSC neurons are derived from patient cells, they represent the complete genetic background of the patients. In the case of *C9orf72* where the repeat is difficult to manipulate genetically, for cloning for example, using the full expansion in the gene context would be ideal. However, iPSC-derived neurons are costly, time- consuming, and only a fraction of the neurons (20-30%) exhibit *C9orf72* pathological characteristics (among others[168,169]).

The recent identification of many new genes, and the failure of many of the mammalian models to fully represent ALS pathogenesis, have driven the use of small animal models, including yeast, zebrafish, flies or worms. Many disease-related genes are highly conserved among species, plus model organisms can be easily genetically manipulated and have rapid reproduction cycles.

The worm *Caenorhabditis elegans* is a multicellular, transparent nematode that is used to study many areas of biology, including aging and stress pathways [170]. Also, with its fully sequenced genome, and its cell lineage fully characterized, it is a highly studied and well characterized model organism. Therefore, *C. elegans* was chosen by our group to model and understand ALS.

LIFESPAN AND GENOME OF C. ELEGANS

C. elegans is a transparent, non-parasitic roundworm. In the wild, the worm is found in soil and includes mainly self-fertilizing hermaphrodites. However, a small proportion of the

population are male worms that can fertilize the hermaphrodite. In laboratory conditions, worms are grown on nematode growth media plate and are fed with *E. coli* [171].

Under normal conditions, the lifecycle of *C. elegans* includes four larval stages (L1-L4) and adulthood. When under harsh environmental conditions during development, such as high temperature, low food supply or high population density, the L1 worm switches to a dauer larval stage instead of the L2 larval stage[172] . The dauer larvae have a distinct morphology and metabolism that allow the worm to survive in difficult conditions[172]. When conditions become favourable, the animal leaves the dauer stage and returns to its normal life cycle[172] . When kept at 20°C, it takes 3-4 days for a wild-type worm to begin laying eggs and it will produce more than 200 during its lifespan. The lifecycle of *C. elegans* is temperature sensitive; kept at 15°C the worms develop slower and temperatures above 25°C can be harmful. At 20°C, a wild-type worm lives on average 20 days. Different signs of aging are visible in *C. elegans* including loss of motility, presence of necrotic cells, presence of oxidized protein, deterioration of different tissues and decline in immune function [173].

The complete genome of *C. elegans* was sequenced in 1998 [174]. It consists of more than 19,000 genes distributed across six chromosomes, whereas 40% of these are found in higher organisms [175]. Many genetic biochemical pathways are highly conserved between *C. elegans* and human including different stress response pathways, the insulin-IGF, apoptosis, necrosis and the innate immune response pathways[176-179] .

C. ELEGANS NERVOUS SYSTEM

The *C. elegans* nervous system includes 302 neurons in hermaphrodite worms. Neurons are classified by their functions, their locations or by the neurotransmitter they express.

Ongoing studies aim to completely characterize the interactome of the *C. elegans* nervous system. *C. elegans* have four types of neurons; motor neurons, sensory neurons, interneurons and polymodal neurons [180].

Two types of motor neurons coordinate the movement of the worms; the gamma-aminobutyric acid (GABA) and the cholinergic neurons. Both types of neurons are located on the dorsal side and innervate muscle cells on the dorsal and ventral side of the worm. Cholinergic neurons are involved in locomotion, egg laying, feeding and male mating (reviewed by Rand [181]). Acetylcholine is synthesized by choline acetyl transferase (CHA-1 protein), loaded in vesicles by vesicular transporter (UNC-17protein) and secreted at the synaptic cleft where it activates the acetylcholine receptors on the post-synaptic cells. Subsequently, it is hydrolyzed to be recycled by acetylcholine esterases (ACHE protein) and re-enter the pre-synaptic cell (CHO-1 protein).

GABAergic neurons are inhibitory neurons involved in locomotion and defecation (for a review see Jorgensen[182]). They are activated by cholinergic neurons. GABA is synthesized by glutamic acid decarboxylase (UNC-25 protein), loaded in vesicles by a vesicular transporter (UNC-47 protein) and will activate the inhibitory GABA receptor (UNC-49 protein) on the post-synaptic cells and causing relaxation of the muscle cells.

The tight coordination between the GABAergic and cholinergic neurons allows for the movement of the worm. Hence, when one side of the worm is contracting due to cholinergic activation, the opposite side is relaxing due to GABA inhibition, therefore causing this smooth sinusoidal movement along the animal body [182](Figure 1.3).

Deletion mutant worms of any of the genes involved in GABA or acetylcholine production or secretion, or pharmacological alteration of these pathways will cause abnormal locomotion of the animals. The mutant worms are often used as control when studying worm's locomotion and the promoter of these genes are used to expressed proteins specifically in the motor neurons both of which are often used in ALS worm research.

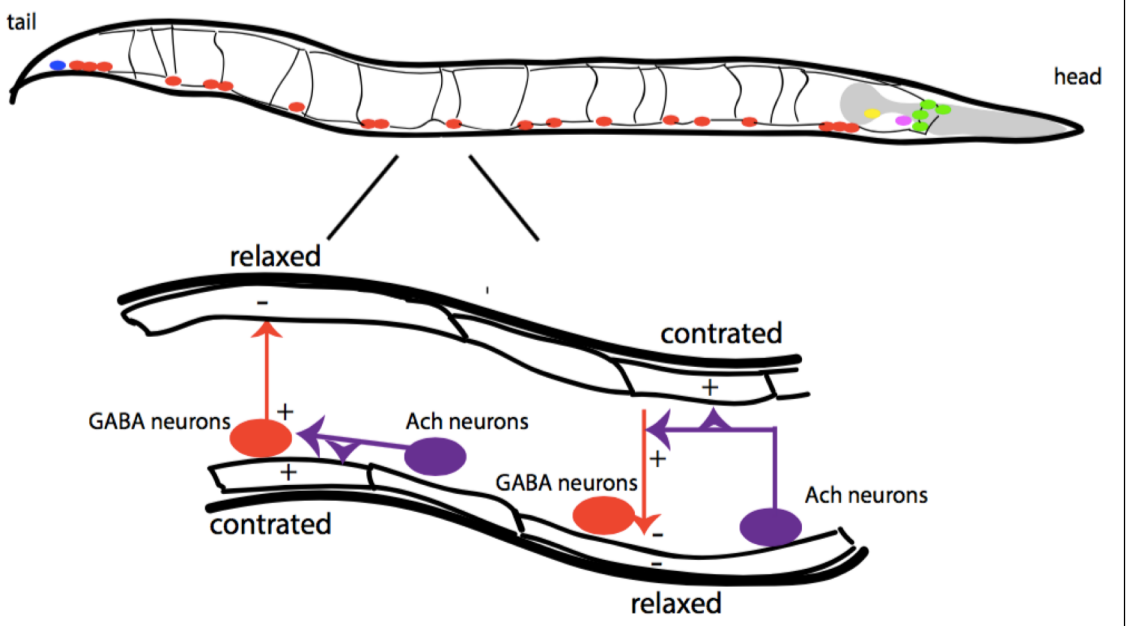


Figure 1.3: *C. elegans* GABAergic motor neurons. Upper: The GABAergic neurons include 20 ventral cord motor neurons innervating either dorsal or ventral muscles, 4 ring motor neurons (green) innervating the head and 3 interneurons (pink, yellow and blue). Bottom: tight coordination between the GABAergic and cholinergic (Ach) motor neurons allow for contraction of one side of the body and relaxation of the opposite side of the animal causing the sinusoidal movement of the animal. Figure inspired by [182]

STRESS RESPONSE PATHWAYS IN C. ELEGANS

Stress response is a major part of *C. elegans* research. Animals can be easily subjected to many different types of environmental stress and many genes involved have turned out to be involved in human diseases such as cancer, infantile diseases and neurodegenerative

diseases [183,184]. The insulin and insulin-like growth factor signaling (IIS) pathway is among the most studied stress response pathways in *C. elegans*. It functions by activating a cell surface receptor, in *C. elegans* DAF-2 (abnormal DAuer Formation 1), that acts to induce a cascade of kinases, including AGE-1 (AGEing alteration1) and AKT-1(AKT kinase family 1), and promotes cell death by inhibiting different transcription factors such as DAF-16 (abnormal DAuer formation 16) and SKN-1(SKiNhead 1) (Figure 1.4). This pathway was shown to also play a role in aging, longevity, fat metabolism, and in different neuronal phenotypes[176]. Mutants that result in reduced IIS are long-lived, stress resistant , exhibit increased fat content and have a reduced neuronal decline.

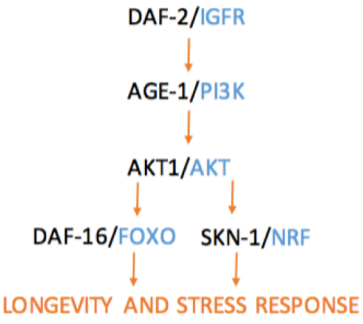


Figure 1.4: The insulin-IGF signaling pathway. In black are the *C. elegans* genes and in blue their human orthologues

The role of the IIS in aging and stress response is mainly regulated by the transcription factor DAF-16. In its unphosphorylated form, DAF-16 is activated and moves to the nucleus where it induces expression of genes associated with longevity and stress resistance. When the IIS is activated, DAF-16 becomes phosphorylated and is retained in the cytoplasm where it cannot affect gene expression [176].

IIS was shown to play an important role in aging of the nervous system. Aged neurons in *C. elegans* display increased branching and neuronal defects, both of which are delayed when the IIS is reduced. Also, DAF-16 is involved in proteostasis and chaperone expression which were also shown to affect neuronal defects. However, not all long-lived mutants acting through DAF-16 affect aging of the nervous system [173]

C. ELEGANS TOOL BOX

Different consortiums and groups have generated many mutant worm strains. Most of these strains are generated by mutagenesis, generating small deletions randomly in the genome. Ethyl methanesulfonate (EMS), and 4,5',8-trimethylpsoralen (TMP) are mainly used for this and deletions or point mutations are generated[185,186]. After mutagenesis, the mutants are then screened for visible phenotypes, such as lethality, progeny numbers, developmental problems, and sequenced to identify the mutations generated. Using this method a large proportion of the genome has been mutagenized. The few genes for which mutations could not be generated are now being inactivated using the most recent genetic tools including the clustered regulatory-interspaced short palindromic repeats/CRISPR associated protein 9 (CRISPR/CAS9) method.

Forward genetics is a valuable tool to identify genes that are involved in a specific function. However, it can be time consuming and the sequencing and identification of the proper mutation that is causing the observed phenotype are the major rate-limiting steps. Therefore, reverse genetic tools have also been developed in *C. elegans.* RNA interference (RNAi) is the process by which one can decrease the expression of a gene by targeting the degradation of its RNA through expressing a complement strand to the target RNA. Most organisms and cells express the machinery to cause gene silencing but it is particularly powerful in *C. elegans* [187]. The worms express a systemic RNAi machinery that includes RNA-directed RNA polymerase that allows for the amplification of the RNA strand, also the RNA strand can be easily expressed in worms [187]. For example, RNA can be expressed in the bacteria that the worms eat and will end up in their intestine. There, it will be transported to different cells across the organism. SID (systemic RNAi defective) proteins are important

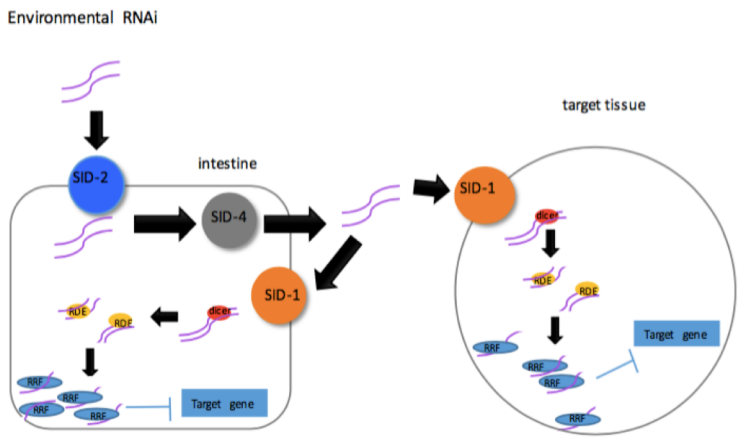


Figure 1.5: RNAi machinery in *C. elegans*. RNA strand in the intestine will be send to all cells of the animal and uptake by SID-1. Then, RNA will be process by DCER, RDF and RRF proteins and bind to its target RNA to degrade it

for export and intake of the RNA. After being taken up by the cell, RNA is cleaved and processed using DCR (Dicer related) and RDE (RNAi defective) proteins. The RNA strand is then amplified by the RRF (RNA-dependant RNA polymerase family) family of proteins and binds to its target mRNA causing its degradation (Figure 1.5). Mutations affecting any of these genes make the worm partially or completely resistant to RNAi.

The RNAi process is highly efficient in almost all cell types with the exception of the nervous system. Neurons do not express the SID proteins, and therefore cannot uptake the RNA strand. However, genetic manipulations have allowed the expression of the SID proteins specifically in neurons resulting in animals sensitized to RNAi in the nervous system [188]. Using similar transgenic and deletion mutant worms, many transgenic worms are available with tissue or cell specific RNAi sensitivity [188-190].

Finally, one of the first uses of green fluorescent protein (GFP) was in *C. elegans* [191]. The worm is transparent, so expression of the GFP protein can be expressed in fusion with a known protein allowing for direct visualization of the protein at different stages of the

worm's life cycle. Translational or transcriptional GFP reporters are also available for many genes to visualize the expression and localization of genes and proteins and the tissue and cellular levels.

The first transgenic worms were created by inserting transgenes in the genome using ultraviolet (UV) or gamma radiation[192]. This caused random integration of the transgene into the genome, often in multiple copies. This technique frequently resulted in copy-number variation and integration site effects rather than specifically looking at a phenotype caused by expression of a sole transgene. Now, new technologies are available and single insertion site methods [193] and CRISPR/Cas9 [194] have led to the generation of new transgenic models with low expression levels of the transgene and targeting the endogenous genes in their genome contexts (promoter, regulatory regions, *etc*.). Hopefully, these new models will better recapitulate what is observed in non-transgenic conditions.

ALS C. ELEGANS MODELS

C. elegans have been previously used to model ALS (Table 1.1) [1,175]. Transgenic models in which neuronal and non-neuronal expression of human SOD1, TDP-43 and FUS proteins in *C. elegans* have been characterized (Table 1.1) [195-203]. The mutant proteins induced an abnormal stress response and protein aggregates in many cases [195-198,200-202]. Transgenic expression of mutant SOD1, TDP-43 and FUS proteins in neurons caused neuronal loss and motility [195,198,199,203].

Being the most studied ALS protein in *C. elegans*, TDP-43 is a good example of the variety of experiments that can be carried in worms to understand its role in neurodegeneration. Many models using various expression patterns of TDP-43 wild-type

or mutant proteins (Δ RNA recognition motif, G290A, A315T, M337V, Δ C' end) have been developed and recapitulate key features observed in patients [195,197-199]. Models expressing TDP-43 in all neurons, or in just a subset of neurons have been characterized [195,197-199]. Mutant TDP-43 proteins generate numerous phenotypes including motility problems, synaptic dysfunctions, protein aggregation, and, neuronal loss [195,197-199]. ER stress response pathways and the innate immune system pathways were both also shown to be involved the toxicity caused by TDP-43[63,204], which recapitulates key features of ALS pathogenesis. Also, the characterization of the worm orthologue of *TARDBP*, *tdp-1*, showed that TDP-1/TDP-43 is involved in regulation of stress response and lifespan pathways [205-207]. By itself, a loss of expression of *tdp-1/TARDBP* can induce motility problem in the worm [206], but it was also shown to alleviate proteotoxicity induced by expression of mutant TDP-43, FUS, SOD1, polyglutamine and progranulin proteins [205,206,208].

Finally, a screen to identify compounds that decrease TDP- 43 aggregation was performed in cell lines and many of the molecules identified were able to suppress the impaired motility phenotype of worms expressing mutant human TDP-43 [209]. Thus, *C. elegans* has proven to be a useful model to better understand TDP-43 toxicity with relevance to ALS.

Most studies using *C. elegans*, have focused on the toxicity of known ALS genes. It is important to note that almost 90% of ALS cases have no known genetic causes. Also, there is wide variation in the onset of symptoms, for ALS, even amongst individuals with the same genetic cause. Furthermore, ALS patients can live between 6 months and 6 years after diagnosis, thus there is speculation that environmental factors may influence disease onset and progression [26]. Several groups have used *C. elegans* to identify compounds that cause

specific degeneration of motor neurons [210-212] opening the door to identifying environmental modifiers of degeneration in ALS models. However, confirmation will be required and large epidemiological studies are needed to evaluate the relevance of these compounds in humans.

OTHER DISEASE RELATED C. ELEGANS MODELS

C. elegans have been widely used to study neurodegenerative disorders. Parkinson's disease, Huntington's disease and Alzheimer's disease models (reviewed in [213-215]) were all developed expressing human mutant proteins in the worms. Motor neuron diseases and non-coding diseases have also been studied.

Hereditary spastic paraplegia (HSP) is a group of diseases affecting mainly the lower motor neurons. This group is genetically heterologous, with more than 40 loci linked to the HSP [216,217]. Using deletion mutants, *spas-1 (SPAStin human neurodegeneration-associated AAA ATPase related 1)* and *nipa-1(nonimprinted gene in Prader/Willi/angelman syndrome region 1 homolog 1)*, the orthologues of human *SPG-4 (Spastic paraplegia 4)* and *NIPA* (*ZC3HC1, zinc finger, C3HC- type containing 1*) respectively, two genes link to HSP, it was shown that these genes have a conserved function between human and *C. elegans* and that deletion of *nipa-1* was detrimental to neurons [218,219] .

Spinal muscular atrophy (SMA) is another motor neuron disorder caused by a decreased expression in *SMN1/2 (Survival in motor neuron proteins 1 and 2*) [220]. A decreased expression of *smn-1(human survival motor neuron gene homolog)*, the

orthologue in *C .elegans*, was shown to induce severe locomotion deficits [221] and known modifiers and interactors of SMN proteins were confirmed in worms [221,222].

Finally, in an attempt to understand more about myotonic dystrophy, RNA expression of CAG or CUG repeats in muscle cells was shown to be toxic and could recapitulate key features of DM1 such as presence of RNA foci [223,224]. Also, similar to what is observed in patients, a decreased expression in *mbl-1(muscleblind splicing regulator homolog)*, the orthologue of human muscle-blind genes, was shown to cause neuromuscular junction abnormalities [225] .

Therefore these models confirm that *C. elegans* can be a suitable model organism to understand more about motor neuron and/or repeat disorders.

RESEARCH HYPOTHESIS

The identification of the non-coding repeat of *C9orf72* as a major player in ALS was first reported at the end of 2011. At the time this project was initiated, in January 2012, very little was known about the cellular impact of this specific repeat. Therefore, the first models were developed to better understand specific phenotypes that were observed in patient cells and tissue, over the years these models were adapted to take into account new findings. Extrapolating about the toxicity of the GGGGCC repeat with other non-coding repeat diseases has been a great value. The main goal of this project was to develop *C. elegans* models to characterized *C9orf72* toxicity and apply our findings to ALS *C9orf72* positive patient cells. More specifically, we have:

Goal #1: To characterize the impact of a decreased expression of *C9orf72* orthologue on the worm nervous system (Chapter 2)

Goal #2: To characterize the impact of the expression of an expanded GGGGCC RNA in *C. elegans* (Chapter 3)

Goal #3: To identify drugs that can alleviate *C9orf72* loss and gain of function toxicity in *C. elegans* and confirm their actions in patient cells (Chapter 3).

CHAPTER 2

INTRODUCTION

The first reports of *C9orf72* involvement in ALS revealed an abnormal expression of *C9orf72* in individuals carrying the pathogenic GGGGCC expansion[47,74]. In patients positive for *C9orf72* expansion, a decreased expression of *C9orf72* mRNA was seen in lymphoblastoid cells, iPSC-derived neurons, frontal cortex, spinal cord and cerebellum [47,74,114,116]. Correlating with a decreased mRNA level, a decreased in protein expression level of *C9orf72* was also observed in motor, temporal and frontal cortexes as well as in the cerebellum and lumbar spinal cord [116,117]. However, a few groups failed to observe this finding in iPSC-derived neurons and frontal cortex [48,226], thus suggesting this reduced expression might vary across different ALS- *C9orf72* patients. Methylation of CpG islands surrounding the repeat and of the GGGGCC repeat itself was speculated to cause the decreased expression. Methylation was observed in ALS and/or FTD patients blood cells; in the hippocampus; frontal, temporal and motor cortices; cerebellum and in spinal cord [227-231].

To understand more about the impact of this decreased expression, we characterized the *alfa-1* gene in *C. elegans*, the sole orthologue of *C9orf72* in the nematode. *alfa-1* (sequence F18A1.6) shares 26% identity and 59% similarity [111,232]. Similar to the human sequence, it encodes two isoforms, a short and a long one with an alternative exon 1, coding for two proteins (731 and 734 amino acids long). Since it is conserved throughout its sequence, it is speculated that even in worms the function of C9ORF72 protein is conserved.

Using a mutant worm carrying a 240 base pair deletion in *alfa-1* gene causing a complete loss of expression, we have evaluated the effect of loss of function of *alfa-1/ C9orf72* in the nematode nervous system.

MANUSCRIPT :*DELETION OF C9ORF72 RESULTS IN MOTOR NEURON DEGENERATION AND STRESS SENSITIVITY IN C. ELEGANS*

PloS One, *8*(12), e83450. http://doi.org/10.1371/journal.pone.0083450.s004

AUTHORS AND AFFILIATIONS

Martine Therrien[1], Guy A. Rouleau[2], Patrick A. Dion[1,3], J. Alex Parker[1,3]

[1]The Centre of Excellence in Neurosciences, CHUM Research Center
[2] Montreal Neurological Institute
[3]Pathology and cell biology department, University of Montreal

AUTHORS CONTRIBUTION

MT and JAP designed the experiments

MT carried the experiments

MT and JAP wrote the manuscript

MT, GAR, PAD, JAP edited the manuscript

MANUSCRIPT

ABSTRACT

An expansion of the hexanucleotide GGGGCC repeat in the first intron of *C9orf72* gene was recently linked to amyotrophic lateral sclerosis. It is not known if the mutation results in a gain of function, a loss of function or if, perhaps both mechanisms are linked to pathogenesis. We generated a genetic model of ALS to explore the biological consequences of a null mutation of the *Caenorhabditis elegans C9orf72* orthologue, *F18A1.6,* also called *alfa-1*. *alfa-1* mutants displayed age-dependent motility defects leading to paralysis and the

specific degeneration of GABAergic motor neurons. *alfa-1* mutants showed differential susceptibility to environmental stress where osmotic stress provoked neurodegeneration. Finally, we observed that the motor defects caused by loss of *alfa-1* were additive with the toxicity caused by mutant TDP-43 proteins, but not by the mutant FUS proteins. These data suggest that a loss of *alfa-1/ C9orf72* expression may contribute to motor neuron degeneration in a pathway associated with other known ALS genes.

INTRODUCTION

Amyotrophic lateral sclerosis (ALS) is one of the most common neurodegenerative disorders and it is characterized by progressive death of motor neurons in the brain and spinal cord. In 1993, the first ALS gene identified was *superoxide dismutase 1* (*SOD1*)[1] and thanks to recent genetic advances there are now over twenty genes linked to ALS [2]. Genes recently shown to be mutated in ALS include the DNA/RNA binding proteins *TAR DNA binding protein 43 (TDP-43)* and *Fused-in-sarcoma (FUS)* [3-6], and *C9orf72*, the latter being a major cause of familial and sporadic ALS [7,8].

GGGGCC repeat expansions are found in the first intron of *C9orf72* and the presence of such long non-coding repeat is suggestive of a toxic gain of function mechanism driving neurodegeneration, perhaps through RNA toxicity, or uncontrolled translation of the repeat into non-native protein species [9]. Very little is known about the biological role of *C9orf72* other than its sequence similarity to the GDP/GTP exchange factor "Differentially Expressed in Normal and Neoplasia" (DENN) [10]. To learn more about the biological role of *C9orf72* we turned to the model organism *Caenorhabditis elegans* and characterized the *C9orf72* orthologue F18A1.6, also called *alfa-1*, in a number of behavioural assays. Although appearing morphologically normal we observed that *alfa-1*(*ok3062*) null mutants developed an age-dependent motor phenotype and neurodegeneration specific to GABAergic motor neurons. Furthermore, *alfa-1(ok3062)* mutants showed hypersensitivity to osmotic stress which further exacerbated motor neuron degeneration. Lastly, we observed that *alfa-1(ok3062)* shows differential genetic interactions with mutant TDP-43 and FUS proteins suggesting a complex interaction amongst some ALS genes.

RESULTS

alfa-1 (ok3062) mutants develop an age-dependent motor phenotype

To better understand the pathogenesis that could result from a decreased expression of *C9orf72*, we examined *ok3062* a null allele of *alfa-1*, the *C. elegans* orthologue of *C9orf72*. ALFA-1 shares 60% homology with C9ORF72 (Blast e-value $2x10^{-15}$) (Figure 1A). The *alfa-1(ok3062)* mutation is a deletion spanning portions of exons 3 and 4 resulting in no detectable *alfa-1* RNA expression (Figure 1B, C). *alfa-1(ok3062)* mutants were superficially normal and had total progeny and lifespan comparable to wild type worms (Figure S1 and Table S2). However, when worms were grown on solid media we observed motility defects when the *alfa-1(ok3062)* mutants reached adulthood, and it ended as an age-dependent paralysis phenotype affecting on average 60% of worms by day 12 of adulthood compared to approximately 20% seen in wild type N2 worms (Figure 2A and Table S1). The progressive paralysis phenotype may indicate impaired transmission at the neuromuscular junction similar to what we previously observed in our ALS models expressing TDP-43 and FUS proteins in *C. elegans* motor neurons [11].

In worms, body movement is coordinated by excitatory input from acetylcholine and inhibitory inputs from GABA [12]. Aldicarb is an acetylcholinesterase inhibitor used to indirectly detect dysfunctional transmission at the neuromuscular junction in *C. elegans* [13], and worms with impaired GABA processing are hypersensitive to aldicarb-induced paralysis [14]. *alfa-1(ok3062)* mutants were more sensitive to aldicarb induced paralysis compared to wild type worms (Figure 2B). These data suggest that *alfa-1(ok3062)* mutants

may have impaired inhibitory GABAergic signalling, perhaps recapitulating the neurotransmitter imbalance observed in ALS patients [15]. When worms are grown in liquid culture they display a swimming behaviour, a vigorous activity that actively engages the neuromuscular junction to maintain activity of the body wall muscles. The paralysis phenotype of *alfa-1(ok3062)* mutants was greatly accelerated when the worms were grown in liquid culture, where approximately 60% of the worms became paralyzed in 8 hours (Figure 2C), compared to 12 days when grown on solid media.

In addition to neuronal dysfunction, our previous TDP-43 and FUS models also showed age-dependent degeneration of motor neurons [11]. Therefore, to assess for similar phenotypes we examined several neuronal populations in our *alfa-1(ok3062)* mutants. To do so, we crossed the *alfa-1(ok3062)* mutation into strains with integrated reporters expressing GFP in different neurons including the GABAergic neurons (*unc-47p::GFP*), the dopaminergic neurons (*dat-1p::GFP*) and cholinergic neurons (*unc-17p::GFP*). At day 9 of adulthood we observed neurodegeneration, in the form of gaps and breaks, only within GABAergic neurons (Figure 2D, E). Thus, our data demonstrate that decreased expression of *alfa-1* causes age-dependent motor defects accompanied by the specific neurodegeneration of the GABAergic motor neurons.

ALFA-1 is required for resistance to osmotic stress

A number of genes linked to ALS have roles in the cellular stress response [16], and *C. elegans* is a convenient system to investigate ALS gene orthologues and stress signaling [17]. To gain further insight into the role of ALS genes and stress, we subjected *alfa-*

1(ok3062) mutants to several, distinct environmental insults. A major regulator of the cellular, and organism-wide stress response signalling in *C. elegans* is the Insulin-IGF pathway. DAF-2 is the sole Insulin/IGF-like receptor in *C. elegans* and hypomorphic mutations in *daf-2* result in stress resistant and long-lived phenotypes compared to wild type animals[18] . In our environmental stress assays, wild type N2 worms are typically stress-sensitive and show progressive lethality while *daf-2(e1370)* animals sre highly resistant to stress-induced lethality. Thus, we asked where *alfa-1(ok3062)* mutants functioned along this stress sensitivity axis. Wild type N2 worms and *alfa-1(ok3062)* mutants were equally sensitive to thermal stress, while *daf-2(e1370)* and *alfa-1(ok3062);daf-2(e1370)* mutants were both highly resistant (Figure 3A). We used the natural compound juglone to test for oxidative stress associated lethality and observed that N2 and *alfa-1(ok3062)* mutants were comparably sensitive, while *daf-2(e1370)* and *alfa-1(ok3062);daf-2(e1370)* mutants were equally resistant to oxidative stress (Figure 3B). Finally, we examined osmotic stress using sodium chloride and observed that *alfa-1(ok3062)* mutants were more sensitive to osmotic stress associated lethality compared to N2 worms. These results were also confirmed by RNAi (Figure S2 A). In the absence of *alfa-1(ok3062), daf-2(e1370)* mutants are slightly less resistant to osmotic stress at a concentration of 400 mM NaCl (Figure 3C). When increasing the concentration to 500 mM NaCl, a significant difference is seen when comparing *alfa-1(ok3062);daf-2(e1370)* to *daf-2(e1370)*, where a loss of *alfa-1* impairs the resistance of daf-2(e1370) worms (Figure 3D). At 600 mM NaCl, both strains die after 60 mins (Figure 3D). *alfa-1(ok3062)* had no effect on dauer formation or the long-lived phenotypes of *daf-2(e1370)* mutants (Figure S2 C, D and

Table S2). These data suggest that ALFA-1 has a specific role in protecting worms against osmotic stress, and might be involve in the Insulin-IGF pathway.

It has been hypothesized that in addition to causative mutations, secondary genetic or environmental factors may contribute to motor neuron degeneration in ALS [19]. Thus, we investigated whether an impaired response to osmotic stress in *alfa-1(ok3062)* worms would impact the degeneration of motor neurons. Using the *unc-47p::GFP* reporter to visualize the GABAergic motor neurons, we subjected *unc-47p::GFP* or *unc-47p::GFP;alfa-1(ok3062)* worms to acute thermal stress which induced comparable levels of neurodegeneration (Figure 3E). However, we observed that acute osmotic stress resulted in a higher rate of motor neurodegeneration in *unc-47p::GFP;alfa-1(ok3062)* animals compared to *unc-47p::GFP* transgenic controls (Figure 3F). The same experiment was carried out using the RNAi hypersensitive strain and similar results were obtained (Figure S2 B) These data suggest that the motor neurons of *alfa-1(ok3062)* animals are specifically sensitive to osmotic stress and that this type of environmental stress may be relevant to the function of C9ORF72.

ALFA-1 differentially interacts with TDP-43 and FUS

There are now over twenty genes linked to ALS and an open question is whether these genes interact to modify neurodegenerative phenotypes. We have previously reported that the neuronal toxicity of dominantly-acting human TDP-43^{A315T} or FUS$^{S57\Delta}$ mutations in *C. elegans* motor neurons can be suppressed by deletion of the worm's TDP-43 orthologue, *tdp-1*[17]. Thus we investigated if *alfa-1* could modify the toxicity of mutant

TDP-43 or FUS proteins in *C. elegans* motor neurons. We generated *TDP-43^{A315T};alfa-1(ok3062)* and *FUS$^{S57\Delta}$;alfa-1(ok3062)* strains and assayed for the age-dependent paralysis phenotype caused by expression of these mutant TDP-43 and FUS proteins. We observed that motor dysfunction was additive for the *TDP-43^{A315T};alfa-1(ok3062)* strain, since the rate of paralysis for this strain was greater than either *alfa-1(ok3062)* or *TDP-43^{A315T}* alone (Figure 4A and Table S1). However, the effects were not additive for *FUS$^{S57\Delta}$;alfa-1(ok3062)* strain since this strain had a comparable rate of paralysis compared to either *alfa-1(ok3062)* or *FUS$^{S57\Delta}$* alone (Figure 4B and Table S1). These data suggest that the genetic interactions between *alfa-1* and TDP-43 or FUS are not equivalent, and that perhaps *alfa-1* and *FUS$^{S57\Delta}$* function in the same pathway, while TDP-43 may use parallel or independent pathways resulting in motor neuron dysfunction.

DISCUSSION

Many questions remain to be answered about the role of *C9orf72* in the pathogenesis of ALS. It is still not clear whether the GGGGCC repeat expansion results in a loss of function, a gain of function or both, or if the size of the repeat has differential effects on these mechanisms of action. Recent reports have observed decreased expression of *C9orf72* when the GGGGCC repeat reaches pathogenic length[7,20,21]. Since no clear mechanisms have been demonstrated for *C9orf72* toxicity, *in vivo* models are important tools to investigate normal biological functions that may lead to insights about the disease state.

C9ORF72 protein sequence is highly similar to ALFA-1 protein sequence. It was hypothesized by two different groups that *C9ORF72* share common feature with DENN proteins[10,22]. Interestingly, Zhang *et al.* have also shown that most amino acids conserved between *C9ORF72* and other DENN proteins are also conserved between *C9ORF72* and ALFA-1[22]. Therefore, we hypothesized that depletion of ALFA-1 represents the depletion of *C9ORF72* and its impact as a DENN protein.

We investigated the biological consequences of deleting the *alfa-1* from *C. elegans* as a putative model for decreased expression of *C9orf72*in ALS. *alfa-1(ok3062)* mutant worms displayed motility defects that progressed into age-dependent paralysis accompanied by the specific neurodegeneration of GABAergic motor neurons. A locomotion deficit caused by a decreased expression of *C9orf72* was recently reported in zebrafish [20] corroborating our results that decreased expression of this protein causes a motor phenotype. Further characterization of ALFA-1 remains to be done, as it will be important to determine that

ALFA-1 is expressed in the nervous system, which has been reported for fish and mouse models [8,20]. However, the rapid onset of motor phenotypes in *alfa-1(ok3062)* mutants in liquid culture sets the stage for chemical screens using neuroprotective molecules. Our nascent drug testing experiments suggest that *alfa-1(ok3062)* toxicity may be distinct from TDP-43, since protective molecules identified previously [23,24]do not suppress motor *alfa-1(ok3062)* phenotypes (data not shown).

Of interest is the recurring theme involving ALS genes in the cellular stress response [16]. We have previously shown that TDP-1, the orthologue of TDP-43 in *C. elegans*, is also involved in response to osmotic stress [17]. Also, in cellular models FUS was shown to robustly react to osmotic stress and increase resistance to this stress [25]. We observed that *alfa-1* mutants were specifically sensitive to osmotic stress and that exposure to this stress enhanced motor neuron degeneration. Cells maintain extensive quality control mechanisms to preserve protein homeostasis against environmental or intrinsic challenges [26]. Osmotic stress can lead to cellular shrinkage, macromolecular crowding and increased protein aggregation with perhaps irreversible degenerative outcomes [27]. Thus it is easy to appreciate the importance of maintaining osmotic balance over the life of a neuron especially since many proteins linked to ALS have a propensity to misfold and aggregate. Here we showed that osmotic stress enhanced neurodegeneration in *alfa-1(ok3062)* mutants, but the impact of osmotic stress on motor neuron health awaits further investigation, which may in time open a new avenue for potential therapeutic strategies.

Recent progress in genetics is developing a more complete picture of the ALS spectrum, but with this information comes the need to better understand the interactions

of ALS genes under normal, pathogenic and aging conditions [2]. Work from genetically expedient model organisms has investigated the genetic interactions of several ALS genes[17,28-30]. However, the picture is not yet complete and nothing is currently known about genetic interactions of *C9orf72* with other ALS genes. We observed that deletion of *alfa-1* enhanced motor defects by mutant TDP-43, but not by mutant FUS. One interpretation of these data is that since the motor neuron degeneration caused by the loss of *alfa-1* is additive to the toxicity of mutant TDP-43 proteins, these mechanisms function in parallel or separate pathways. Oppositely, a mechanism may be shared by *alfa-1* and mutant FUS, as the level of paralysis observed was comparable to each condition alone or in combination. Additional tools and experiments are required to better understand the basis of these genetic interactions.

The molecular pathogenic mechanisms behind the genetic mutations of many ALS genes are not fully understood. For many of these genes it is not known whether mutations lead to a gain of function, a loss of function, or both. An informative example comes from studies of TDP-43 where several *in vivo* models suggested that both occur simultaneously [17,31,32]. A similar situation may exist for *C9orf72* where loss of expression leads to motor phenotypes, in conjunction with recent findings demonstrating that the expression of GGGGCC RNA is toxic in a *Drosophila* model [33], and that repeats can be inappropriately translated into different peptides with additional potential cytotoxic effects [34,35]. Thus, further characterization of both mechanisms will unravel the toxicity caused by the presence of the GGGGCC repeat in the first intron of *C9orf72*.

MATERIALS AND METHODS

Nematode strains

Standard methods for culturing and handling the worms were used (Stiernagle, 2006). N2, *daf-2(e1370)*, *F18A1.6(ok3062), oxIs12 [unc-47p::GFP + lin-15(+)]*, *vsIs48 [unc-17::GFP], vtIs1 [dat-1p::GFP;rol-6(su1006)], rrf-3(pk1426)* were obtained form the *C. elegans* Genetics Center (University of Minnesota, Minneanapolis) and maintained at 20°C on standard NMG petri streaked with OP-50 *E. coli*. Transgenic $FUS^{S57\Delta}$ and $TDP\text{-}43^{A315T}$ strains were previously described (Vaccaro, Tauffenberger, Aggad, et al., 2012b). *F18A1.6(ok3062)* was outcrossed to wild type N2 five times before use.

RT-PCR

RNA was extracted using Trizol. After worm lysis and homogenization, chloroform was added and tubes were centrifuged. RNA was precipitated from the aqueous phase using isopropanol, pellets were washed with 75% ethanol and resuspended in water. RNA was reverse transcribed with the QuantiTect kit (Qiagen) preceded by gDNA wipeout. 1 µl of cDNA was used for *act-3* and F18A1.6 amplification using the following primers; F18A1.6 forward 5' AATGAGCGGAACATCAAGC 3', F18A1.6 reverse 5' TTCGGATATGTCAGGCTGAAG 3', *act-3* forward 5'GTTGCCGCTCTTGTTGTAGAC 3', *act-3* reverse 5' GGAGAGGACAGCTTGGATGG3'

Paralysis assay

30 adult, day one worms were transferred to NGM with FUDR plates and scored daily for movement. Worms were counted as paralysed if they failed to move after being prodded in the nose. Experiments were conducted at 20°C and done in triplicates. Survival curves were produced and compared using the Log-rank (Mantel-Cox) test.

Aldicarb test

30 adult, day one worms were transferred to NGM with 1mM aldicarb plates. Worms were scored every 30 minutes for two hours and counted paralysed if they failed to move after being prodded on the nose. Experiments were conducted at 20°C and done in triplicates. Survival curves were produced and compared using the Log-rank (Mantel-Cox) test using GraphPad Prism software.

Liquid culture

Synchronized populations of worms was obtained by hypochlorite extraction. 20-30 young adults were distributed in 96-well plate containing OP50 and incubated for eight hours at 25°C.Worms were counted paralysed if they failed to moved after gently tapping the side of the plate. The mean and SEM were calculated and two-tailed t-tests were used for statistical analysis.

Lifespan assay

30 adult, day one worms were transferred to NGM-FUDR plates and counted every two days. Worms were counted as dead if they did not respond to tactile stimulus. Survival curves were produced and compared using the Log-rank (Mantel-Cox) test.

RNAi experiments

RNAi-treated strains were fed with *E. coli* (HT115) containing an empty vector (EV) or *alfa-1* (F18A1.6) RNAi clones from the ORFeome RNAi librairy (Open Biosystems). RNAi experiments were performed at 20°C. Worms were grown on NGM enriched with 1mM Isopropyl-b-thiogalacto-pyranoside (IPTG).

Stress assays

Worms were grown at 20°C on normal NGM plates until day one of adulthood. 30 adult, day one worms were then transferred to NGM plates + 240 μM juglone (oxidative stress), or NGM + 400 mM NaCl, or NGM + 500mM NaCl, or NGN + 600 mM NaCl (osmotic stress). Tests were carried at 20°C for oxidative and osmotic stresses and at 37°C for thermal stress. Worms were counted every two hours for up to 14 hours. For all experiments, worms were counted as dead if they did not respond to tactile stimulus. Survival curves were produced and compared using the Log-rank (Mantel-Cox) test.

Neurodegeneration assay

To score gaps or breaks, synchronized animals were selected at day one, five and nine of adulthood for *in vivo* visualization. For neurodegeneration count during stress tests, adult day one worms were transferred to NGM + 400 mM NaCl at 20°C (osmotic stress) or normal NGM and put at 37°C (thermal stress) for six hours. To confirm those results with RNAi, *rrf-3(pk1426)* worms submitted to *alfa-1* or EV RNAi up to day 1 of adulthood. Worms were then transferred on 400 mM NaCl for six hours. For visualization, animals were immobilized in M9 with 5 mM of levamisole and mounted on slides with 2% agarose pads. Neurons were visualized with a Leica 6000 microscope and a Leica DFC 480 camera. For all experiments, a minimum of 100 worms was scored over at least 3 trials for all conditions. The mean and SEM were calculated and two-tailed t-tests were used for statistical analysis.

ACKNOWLEDGMENTS

J.A.P is a CIHR New Investigator. This research was supported by the CHUM Foundation, ALS Canada, and the Muscular Dystrophy Association (USA).

REFERENCES

1. Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, et al. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature. 1993;362: 59–62. doi:10.1038/362059a0
2. Turner MR, Hardiman O, Benatar M, Brooks BR, Chiò A, de Carvalho, et al. Controversies and priorities in amyotrophic lateral sclerosis. The Lancet Neurology. Elsevier Ltd; 2013;12: 310–322. doi:10.1016/S1474-4422(13)70036-X
3. Kabashi E, Valdmanis PN, Dion P, Spiegelman D, McConkey BJ, Vande Velde C, et al. TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. Nat Genet. 2008;40: 572–574. doi:10.1038/ng.132
4. Kwiatkowski TJ, Bosco DA, LeClerc AL, Tamrazian E, Vanderburg CR, Russ C, et al. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. Science. 2009;323: 1205–1208. doi:10.1126/science.1166066
5. Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, et al. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science. 2008;319: 1668–1672. doi:10.1126/science.1154584
6. Vance C, Rogelj B, Hortobágyi T, De Vos KJ, Nishimura AL, Sreedharan J, et al. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science. 2009;323: 1208–1211. doi:10.1126/science.1165942
7. DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, et al. Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. Neuron. 2011;72: 245–256. doi:10.1016/j.neuron.2011.09.011
8. Renton AE, Majounie E, Waite A, Simón-Sánchez J, Rollinson S, Gibbs JR, et al. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. Neuron. 2011;72: 257–268. doi:10.1016/j.neuron.2011.09.010
9. Cruts M, Gijselinck I, Van Langenhove T, van der Zee J, Van Broeckhoven C. Current insights into the C9orf72 repeat expansion diseases of the FTLD/ALS spectrum. Trends Neurosci. 2013;36: 450–459. doi:10.1016/j.tins.2013.04.010
10. Levine TP, Daniels RD, Gatta AT, Wong LH, Hayes MJ. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. Bioinformatics. 2013;29: 499–503. doi:10.1093/bioinformatics/bts725
11. Vaccaro A, Tauffenberger A, Aggad D, Rouleau G, Drapeau P, Parker JA. Mutant TDP-43 and FUS cause age-dependent paralysis and neurodegeneration in C. elegans. PLoS ONE. 2012;7: e31321. doi:10.1371/journal.pone.0031321
12. Jorgensen EM. GABA. WormBook. 2005;: 1–13. doi:10.1895/wormbook.1.14.1
13. Mahoney TR, Luo S, Nonet ML. Analysis of synaptic transmission in Caenorhabditis

elegans using an aldicarb-sensitivity assay. Nat Protoc. 2006;1: 1772–1777. doi:10.1038/nprot.2006.281
14. Vashlishan AB, Madison JM, Dybbs M, Bai J, Sieburth D, Ch'ng Q, et al. An RNAi screen identifies genes that regulate GABA synapses. Neuron. 2008;58: 346–361. doi:10.1016/j.neuron.2008.02.019
15. Foerster BR, Pomper MG, Callaghan BC, Petrou M, Edden RAE, Mohamed MA, et al. An imbalance between excitatory and inhibitory neurotransmitters in amyotrophic lateral sclerosis revealed by use of 3-T proton magnetic resonance spectroscopy. JAMA Neurol. 2013;70: 1009–1016. doi:10.1001/jamaneurol.2013.234
16. Li YR, King OD, Shorter J, Gitler AD. Stress granules as crucibles of ALS pathogenesis. The Journal of Cell Biology. Rockefeller Univ Press; 2013;201: 361–372. doi:10.1083/jcb.201302044
17. Vaccaro A, Tauffenberger A, Ash PEA, Carlomagno Y, Petrucelli L, Parker JA. TDP-1/TDP-43 regulates stress signaling and age-dependent proteotoxicity in Caenorhabditis elegans. PLoS Genet. Public Library of Science; 2012;8: e1002806. doi:10.1371/journal.pgen.1002806
18. Kenyon CJ. The genetics of ageing. Nature. 2010;464: 504–512. doi:10.1038/nature08980
19. Polymenidou M, Cleveland DW. The Seeds of Neurodegeneration: Prion-like Spreading in ALS. Cell. Elsevier Inc; 2011;147: 498–508. doi:10.1016/j.cell.2011.10.011
20. Ciura S, Lattante S, Le Ber I, Latouche M, Tostivint H, Brice A, et al. Loss of function of C9orf72 causes motor deficits in a zebrafish model of Amyotrophic Lateral Sclerosis. Ann Neurol. 2013;: 180–187. doi:10.1002/ana.23946
21. Xi Z, Zinman L, Moreno D, Schymick J, Liang Y, Sato C, et al. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. Am J Hum Genet. 2013;92: 981–989. doi:10.1016/j.ajhg.2013.04.017
22. Zhang D, Iyer LM, He F, Aravind L. Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. Front Genet. 2012;3: 283. doi:10.3389/fgene.2012.00283
23. Vaccaro A, Patten SA, Aggad D, Julien C, Maios C, Kabashi E, et al. Pharmacological reduction of ER stress protects against TDP-43 neuronal toxicity in vivo. Neurobiol Dis. 2013;55: 64–75. doi:10.1016/j.nbd.2013.03.015
24. Vaccaro A, Patten SA, Ciura S, Maios C, Therrien M, Drapeau P, et al. Methylene blue protects against TDP-43 and FUS neuronal toxicity in C. elegans and D. rerio. PLoS ONE. 2012;7: e42117. doi:10.1371/journal.pone.0042117
25. Sama RRK, Ward CL, Bosco DA. Functions of FUS/TLS From DNA Repair to Stress Response: Implications for ALS. ASN Neuro. SAGE Publications; 2014;6: 1759091414544472. doi:10.1177/1759091414544472
26. Hoppe T. Too salty for worms: hypertonic stress challenges proteostasis networks. Focus on "Hypertonic stress induces rapid and widespread protein damage in C. elegans". Am J Physiol, Cell Physiol. 2011;301: C555–6. doi:10.1152/ajpcell.00206.2011
27. Burkewitz K, Choe KP, Choung-Hee Lee E, Deonarine A, Strange K. Characterization of the Proteostasis Roles of Glycerol Accumulation, Protein Degradation and Protein Synthesis during Osmotic Stress in C. elegans. Gasset M, editor. PLoS ONE. 2012;7:

e34153. doi:10.1371/journal.pone.0034153
28. Kabashi E, Bercier V, Lissouba A, Liao M, Brustein E, Rouleau GA, et al. FUS and TARDBP but not SOD1 interact in genetic models of amyotrophic lateral sclerosis. PLoS Genet. 2011;7: e1002214. doi:10.1371/journal.pgen.1002214
29. Wang J, Farr GW, Hall DH, Li F, Furtak K, Dreier L, et al. An ALS-Linked Mutant SOD1 Produces a Locomotor Defect Associated with Aggregation and Synaptic Dysfunction When Expressed in Neurons of Caenorhabditis elegans. PLoS Genet. 2009;5: e1000350. doi:10.1371/journal.pgen.1000350.s015
30. Zhang T, Hwang HY, Hao H, Talbot C, Wang J. Caenorhabditis elegans RNA-processing Protein TDP-1 Regulates Protein Homeostasis and Life Span. Journal of Biological Chemistry. 2012;287: 8371–8382. doi:10.1074/jbc.M111.311977
31. Kabashi E, Lin L, Tradewell ML, Dion PA, Bercier V, Bourgouin P, et al. Gain and loss of function of ALS-related mutations of TARDBP (TDP-43) cause motor deficits in vivo. Hum Mol Genet. Oxford University Press; 2010;19: 671–683. doi:10.1093/hmg/ddp534
32. Wang J-W, Brent JR, Tomlinson A, Shneider NA, McCabe BD. The ALS-associated proteins FUS and TDP-43 function together to affect Drosophila locomotion and life span. J Clin Invest. 2011;121: 4118–4126. doi:10.1172/JCI57883
33. Xu Z, Poidevin M, Li X, Li Y, Shu L, Nelson DL, et al. Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. Proceedings of the National Academy of Sciences. 2013;110: 7778–7783. doi:10.1073/pnas.1219643110
34. Mori K, Weng S-M, Arzberger T, May S, Rentzsch K, Kremmer E, et al. The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. Science. American Association for the Advancement of Science; 2013;339: 1335–1338. doi:10.1126/science.1232927
35. Ash PEA, Bieniek KF, Gendron TF, Caulfield T, Lin W-L, DeJesus-Hernandez M, et al. Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. Neuron. 2013;77: 639–646. doi:10.1016/j.neuron.2013.02.004

FIGURES

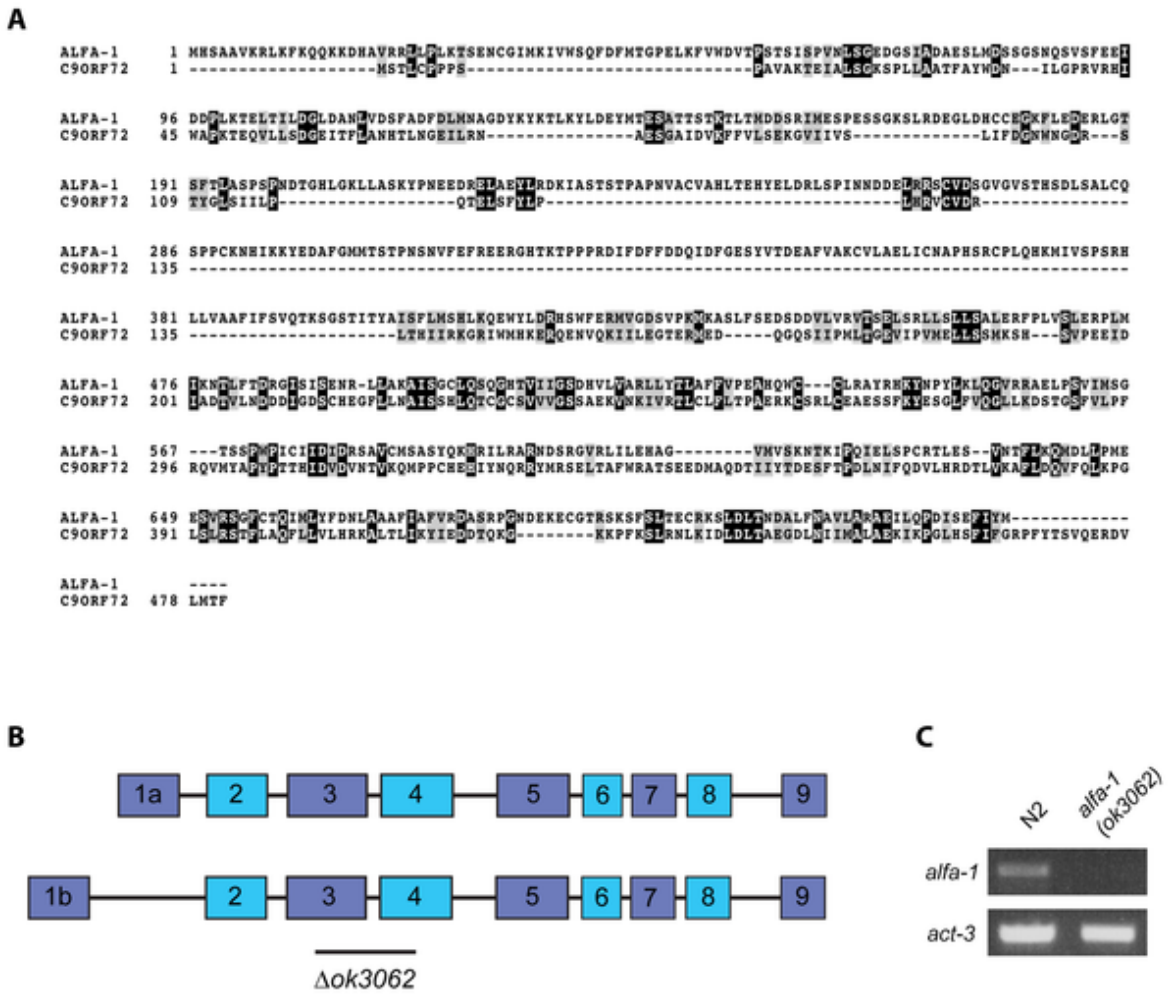


Figure 1

Figure 2.1. *alfa-1* is the orthologue of *C9orf72* in *C. elegans*.
(A) Protein sequence alignment (using Clustal W) of *C9orf72* isoform 1 and ALFA-1 isoform 1. Overall, these sequences share 26% identify and 59% similarity.
(B)*alfa-1* has two predicted transcripts and the *ok3062* deletion mutation spans exons 3 and 4 for both transcripts.
(C) RT-PCR confirming the complete loss of expression of the *alfa-1* transcripts. *act-3* was used as a control.

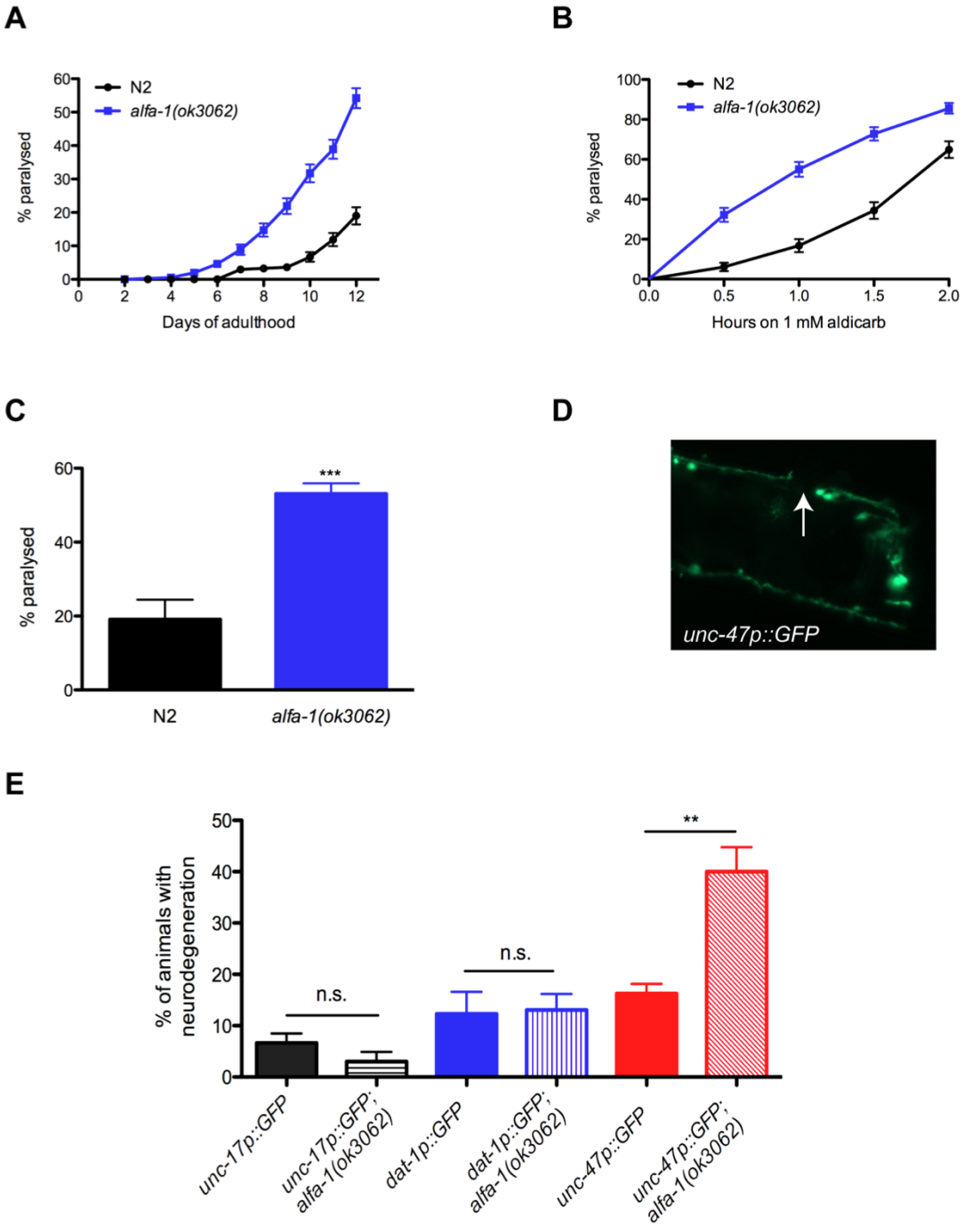

A
N2
alfa-1(ok3062)
% paralysed
Days of adulthood
B
N2
alfa-1(ok3062)
% paralysed
Hours on 1 mM aldicarb
C
% paralysed
***
N2
alfa-1(ok3062)
D
unc-47p::GFP
E
% of animals with neurodegeneration
n.s.
n.s.
**
unc-17p::GFP
unc-17p::GFP; alfa-1(ok3062)
dat-1p::GFP
dat-1p::GFP; alfa-1(ok3062)
unc-47p::GFP
unc-47p::GFP; alfa-1(ok3062)

Figure 2

Figure 2.2. Age-dependent motility defects and neurodegeneration in *alfa-1(ok3062)* mutants.
(A) *alfa-1(ok3062)* mutants showed motility defects leading to paralysis of 60% of the population by day 12 of adulthood compared to 20% for N2 worms (P<0.0001).
(B) *alfa-1(ok3062)* worms are more sensitive to aldicarb-induced paralysis than N2 worms (P<0.0001).
(C) Percentage of wild type N2 or *alfa-1(ok3062)* worms displaying a swimming-induced paralysis phenotype after 8 hours in liquid culture (** P<0.001).
(D) Example of gap (indicated by arrow) along a neuronal process in animals expressing the *unc-47p::GFP* reporter.
(E) Quantification of neurodegeneration at day 9 of adulthood associated with *alfa-1(ok3062)* in different neuronal populations including cholinergic neurons marked by *unc-17p::GFP*, dopaminergic neurons visualized with *dat-1p::GFP*, or GABAergic neurons revealed by *unc-47p::GFP*. Significant neurodegeneration was observed in the GABAergic neurons of *alfa-1(ok3062)* mutants (**P<0.001).

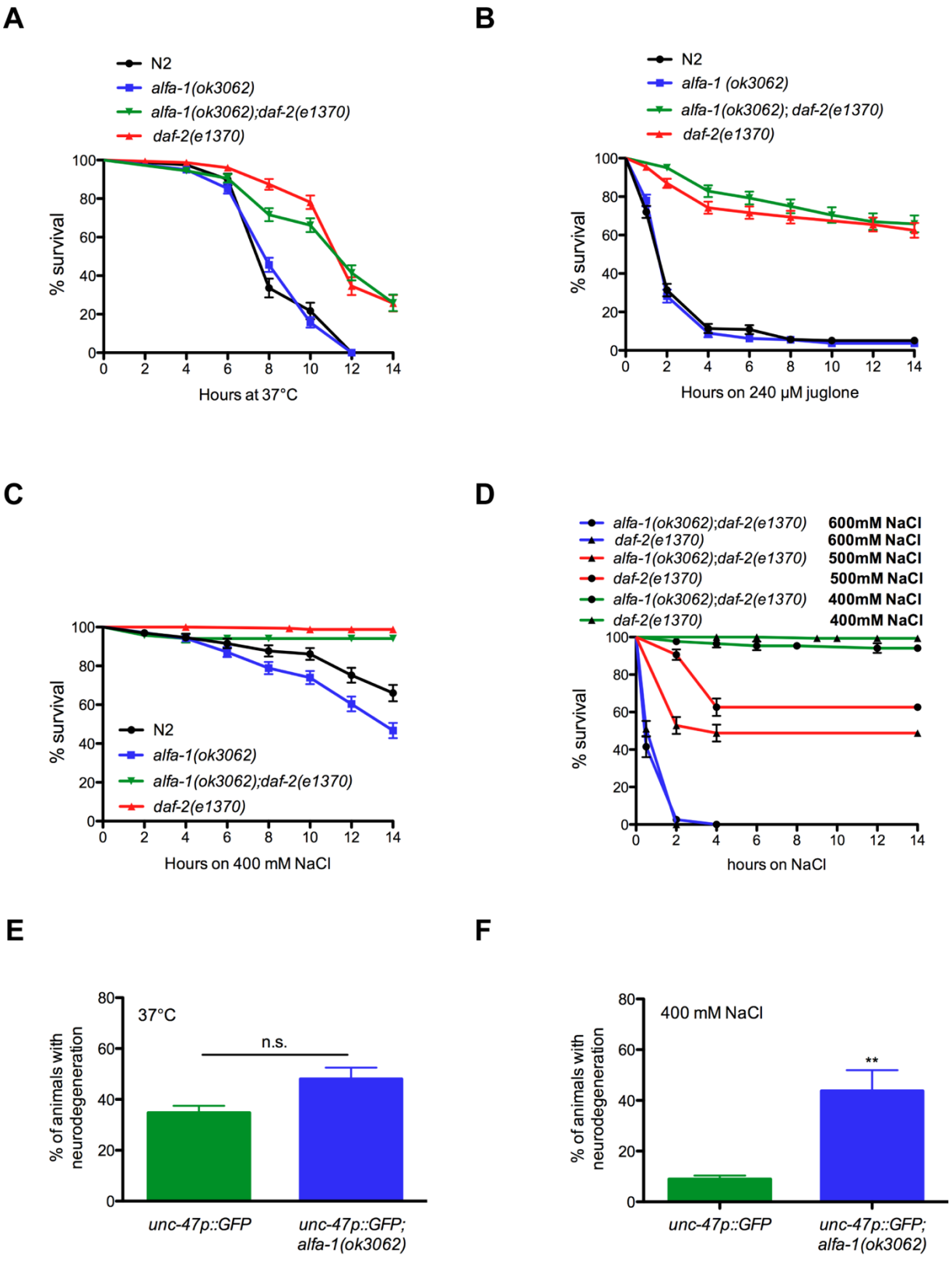

A
N2
alfa-1(ok3062)
alfa-1(ok3062);daf-2(e1370)
daf-2(e1370)
% survival
Hours at 37°C
B
N2
alfa-1 (ok3062)
alfa-1(ok3062); daf-2(e1370)
daf-2(e1370)
% survival
Hours on 240 µM juglone
C
% survival
N2
alfa-1(ok3062)
alfa-1(ok3062);daf-2(e1370)
daf-2(e1370)
Hours on 400 mM NaCl
D
alfa-1(ok3062);daf-2(e1370) 600mM NaCl
daf-2(e1370) 600mM NaCl
alfa-1(ok3062);daf-2(e1370) 500mM NaCl
daf-2(e1370) 500mM NaCl
alfa-1(ok3062);daf-2(e1370) 400mM NaCl
daf-2(e1370) 400mM NaCl
% survival
hours on NaCl
E
37°C
n.s.
% of animals with neurodegeneration
unc-47p::GFP
unc-47p::GFP; alfa-1(ok3062)
F
400 mM NaCl
**
% of animals with neurodegeneration
unc-47p::GFP
unc-47p::GFP; alfa-1(ok3062)

Figure 3

Figure 2.3. *alfa-1(ok3062)* mutants are sensitive to osmotic stress.
(A) In thermal stress resistance assays, *alfa-1(ok3062)* mutants were indistinguishable from wild type N2 worms, and *daf-2*(e1370) were not statistically different *alfa-1(ok3062)*;*daf-2*(e1370) mutants.
(B) In oxidative stress resistance assays, *alfa-1(ok3062)* mutants were indistinguishable from wild type N2 worms, and *daf-2*(e1370) were not statistically different *alfa-1(ok3062)*;*daf-2*(e1370) mutants.
(C) *alfa-1(ok3062)* mutants were more sensitive to osmotic stress than N2 worms ($P<0.005$), while *alfa-1(ok3062)*;*daf-2*(e1370) worms are slightly more sensitive when compared to *daf-2*(e1370) worms alone.
(D) The difference in sensitivity between *alfa-1(ok3062)*;*daf-2*(e1370) and *daf-2*(e1370) increases at 500 mM NaCl ($P<0.005$). At 600 mM NaCl, the effect of NaCl is too drastic to see a difference.
(E) During thermal stress, *alfa-1(ok3062)* worms are not more sensitive to neurodegeneration than the *unc-47p::GFP* worms.
(F) When exposed to 400 mM NaCl, *alfa-1(ok3062)* worms had a higher rate of neurodegeneration than *unc-47p::GFP* worms (*$P<0.05$).

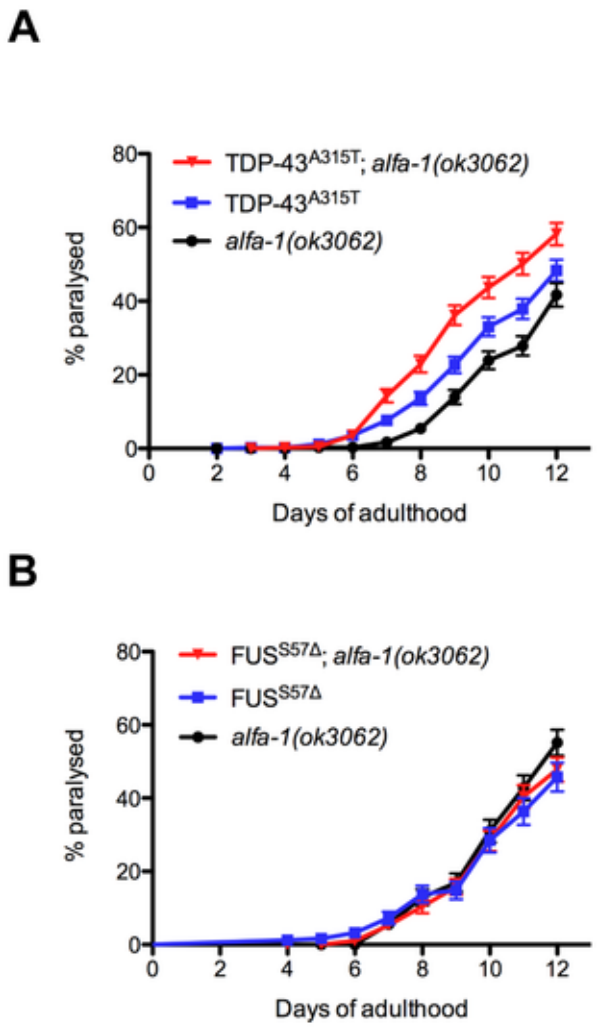


Figure 4

Figure 2.4. Genetic interactions between *alfa-1(ok3062),* TDP-43, and FUS.
(A) *TDP-43^{A315T}*; *alfa-1(ok3062)* worms had a higher rate of paralysis that either *TDP-43^{A315T}* or *alfa-1(ok3062)* worms alone (P<0.005).
(B) *FUS$^{S57\Delta}$* worms , *alfa-1(ok3062)* worms, and *FUS$^{S57\Delta}$; alfa-1(ok3062)* worms showed similar rates of paralysis.

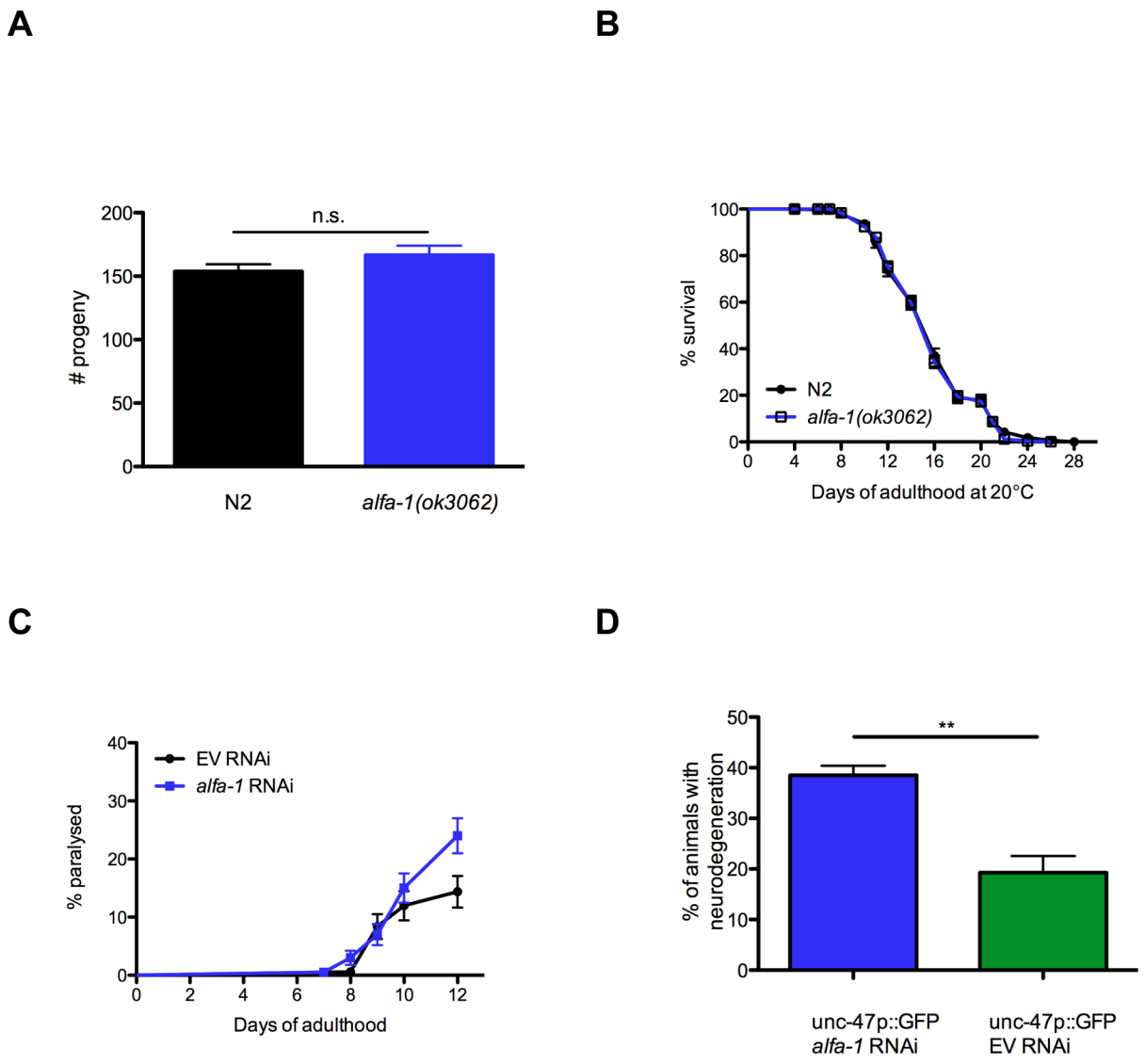


Figure S1

Figure 2.S1

(A)(B) *alfa-1(ok3062)* worms had normal (A) progeny and (B) lifespan compared to N2 worms.

(C) *rrf-3(pk1426)* worms submitted to *alfa-1* RNAi display motility defects causing paralysis at day 12 of adulthood compared to *rrf-3(pk1426)* worms submitted to empty vector (EV).

(D) *rrf-3(pk1426)* worms submitted to *alfa-1* RNAi have increased neurodegeneration at day 9 of adulthood compared to *rrf-3(pk1426)* worms submitted to EV RNAi.

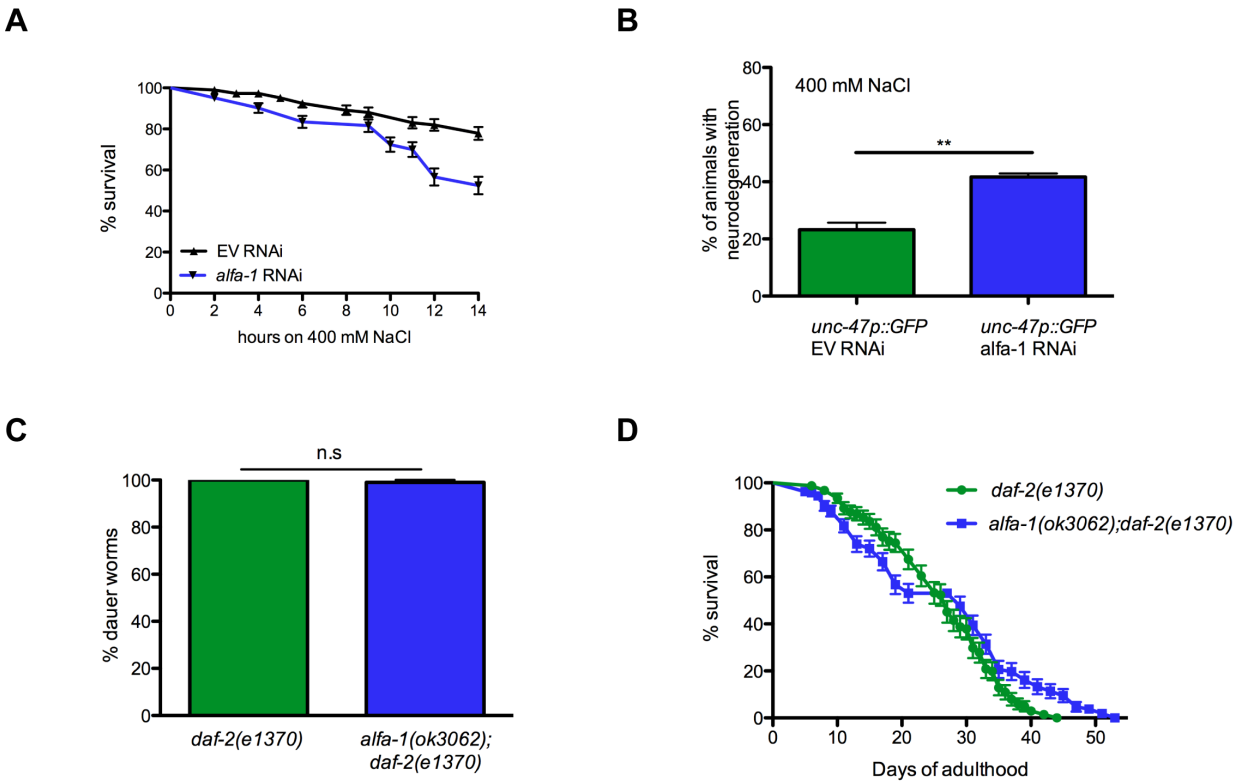


Figure S2

Figure 2.S2
(A) N2 worms subjected to RNAi against *alfa-1* are more sensitive to 400 mM NaCl than those submitted to empty vector (EV) (P<0.0001)
(B) *rrf-3(pk1426)* worms submitted to *alfa-1* RNAi showed an increase neurodegenration of GABAergic motor neurons (*unc47p::GFP*) after 6 hours under osmotic stress compared to *rrf-3(pk1426)* worms submitted EV in the same conditions (P<0.0001).
(C) (D) The *alfa-1(ok3062)* mutation had no effect on (C) dauer formation or (D) the extended lifespan of *daf-2(e1370)* mutants.

	Strain	P-value	Number animals Paralysed/Total
Figure 1A	N2		47/382
	alfa-1(ok3062)	<0.0001	158/375
Figure 4A	TDP-43^{A315T}		153/438
	alfa-1(ok3062)	0.009	111/389
	TDP-43^{A315T};*alfa-1(ok3062)*	0.001	172/397
Figure 4B	FUS$^{S57\Delta}$		80/249
	alfa-1(ok3062)	ns 0.17	116/255
	FUS$^{S57\Delta}$;*alfa-1(ok3062)*	ns 0.79	116/296

Table 2.S1. Paralysis tests for all experiments, ns=non significant.

	Strain	P-value	Number animals Dead/Total
Figure S1B	N2		235/400
	alfa-1 (ok3062)	ns 0.68	305/414
Figure S2B	*alfa-1 (ok3062);daf-2(e1370)*	ns 0.1184	137/216
	daf-2(e1370)		65/134

Table 2.S2. Lifespan assay for all experiments, ns=non significant.

CHAPTER 3

INTRODUCTION

Even though the loss of expression of *C9orf72* was detrimental to neurons in *C. elegans* we could not underestimate the toxicity that might be caused by the presence of an expanded GGGGCC repeat in *C9orf72*. Different teams have speculated about a possible toxic gain of function of the RNA containing an expanded GGGGCC repeat. The repeat might underlie similar event to those observed for other expanded non-coding repeat expansion, like for instance the formation of RNA foci and the production of peptides in a non-ATG manner (RAN translation).

The presence of the expanded GGGGCC repeat in the *C9orf72* V2 and V3 isoforms is speculated to affect the structure of the RNA. Analysis of the GGGGCC repeat by itself has shown that the G/C rich RNA forms a G-quadruplex [233] which could cause abnormal protein and RNA interactions and repeat instability[234]. Molecules affecting the RNA structure of G/C rich repeat, such as TMPγP4, were shown to also change the structure of the GGGGCC repeat of *C9orf72* [234].

In patient tissues, the presence of RNA foci caused by the accumulation of sense and anti-sense *C9orf72* transcripts containing the GGGGCC repeat was observed in the nucleus of neuronal cells [47]. Also, even though the repeat is located in the intron, it can initiate translation in a non-ATG manner, leading to the production of potentially toxic dipeptides repeat proteins [235,236]. In *C9orf72*, sense and antisense transcripts of *C9orf2* lead to the production of glycine-alanine, glycine-proline and glycine-arginine peptides from the sense

transcript and proline-arginine, glycine-proline and proline-alanine peptide from the antisense transcript

Abnormal RNA processing was shown to be a major player in ALS pathogenesis [237]. *C9orf72* transcripts were shown to cause the formation of RNA and protein aggregates and to be involved in stress granule formation all of which, could impair RNA metabolism as well. Therefore, many groups have tried to evaluate the impact of *C9orf72* on global RNA expression by whole-transcriptome sequencing [114,226,238]. Even though no RNA was clearly identified as a consequence of the presence of *C9orf72*, this method is useful to evaluate global RNA metabolism [114,226,238].

The small size of *C. elegans*, its rapid life cycle, its ease of cultivation and ability to obtain large numbers of animals make it an attractive model for drug discovery. Furthermore, worms can be grown in liquid culture which is easy to adapt for drug screening purposes [239]. Boyd *et al.* , using a TDP-43 trangenic model, have shown that drugs identified from cells often have relevance in *C. elegans* [209].

In this chapter, we have developed a new model to understand the toxicity caused by an expanded GGGGCC repeat RNA. Using this new *C. elegans* model and the one developed in the previous chapter (*alfa-1* loss of function), a drug screen was carried where more than 4,000 compounds were assessed for their effect on *alfa-1*/ *C9orf72* toxicity and confirmed in our RNA GGGGCC toxicity model. Finally, the availability of cells from *C9orf72* positive patients has allowed us to confirm the effect of those drugs in a mammalian cell model.

MANUSCRIPT: DRUG SCREENING IDENTIFIES MOLECULES THAT CAN ALLEVIATE THE TOXICITY CAUSED BY LOSS OF *C9ORF72* AND BY TOXIC GGGGCC REPEAT EXPANSION RNA IN *IN VITRO* AND IN *VIVO MODELS*

Authors:

Martine Therrien[1,2], Simon Girard[3], Jame Doyle[1,4], Amélie Quoibion[4], Claudia Maios[1], Guy A. Rouleau [3,4], Patrick A. Dion[3,4], J. Alex Parker[1,2]

Affiliations:

[1] CHUM Research Center
[2] University of Montreal
[3] McGill University
[4] Montreal Neurological Institute

Authors contribution

MT, PAD and JAP designed the experiments

MT, SG, JD, AM and CM carried the experiments

MT and JAP wrote the manuscript

MT, GAR, PAD, JAP edited the manuscript

ABSTRACT

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disorder affecting upper and lower motor neurons. 10% of the cases are familial and the GGGGCC expansion in *C9orf72* gene is the most common known cause of ALS thus far. Many researchers have attempted to explain the toxicity of the pathogenic GGGGCC repeat, including loss and gain of function mechanisms. We show here, that in the nematode *C. elegans*, gain of function of the GGGGCC repeat and loss of *alfa-1* are toxic to neurons. A high-throughput screen was carried to identify molecules that could alleviate both types of toxicity in worms and identified eight molecules that could be neuroprotective. From these, two also restore that abnormal gene expression observed in *C9orf72* positive patient-derived cell lines.

INTRODUCTION

Amyotrophic lateral sclerosis is a fatal, neurodegenerative disorder. It causes progressive loss of the motor neurons leading to muscle weakness. Death typically occurs 3-5 years after diagnosis from respiratory failure due to denervation of the respiratory muscles[1]. High variability is observed among patients regarding disease age of onset and progression. Even though more than 15 molecules have been tested in phase 2-3 clinical trials over the past decade[2], riluzole is still the only drug approved to treat ALS patients[1]. Therefore, there is an urgent need to identify new, effective therapeutic molecules for ALS patients.

In most cases of ALS, there is no familial history of the disease, and those cases are referred to as sporadic ALS (sALS), but 10% of all ALS cases do and are referred to as familial ALS (fALS)[3]. More than 20 genes are genetically linked to ALS, including *SOD1*, *TARDBP*, *FUS* and *C9orf72*[4]. Being the cause of more than 30% of fALS cases and of 7% of sALS cases, the expansion of GGGGCC repeat in the first intron of *C9orf72* is the most prevalent known cause of ALS thus far identified[5]. The presence of more than 30 units of the GGGGCC repeat was shown to cause a decreased expression of *C9orf72*, formation of nuclear RNA foci and production of different dipeptides repeats through abnormal translation mechanisms [6]. The mechanism by which the presence of the expanded GGGGCC repeat is toxic to motor neurons is still unclear.

To better understand the toxicity of *C9orf72,* and to identify molecules that could alleviate it, we turned to the worm *Caenorhabditis elegans*. Using this model organism, we show that both expanded GGGGCC repeat RNA, and loss of *alfa-1*, the sole orthologue of *C9orf72*, is toxic to the animal's motor neurons. Using these models, we conducted a drug

screen and identified molecules that reduce both forms of toxicity. Two of those drugs were also shown to ameliorate abnormal RNA expression in *C9orf72* positive cell lines, suggesting they may be useful for testing in clinical settings.

RESULTS

Generation of a *C. elegans* GGGGCC$_{50}$ RNA model

To evaluate the toxicity caused by the expression of a pathogenic GGGGCC repeat of *C9orf72* we generated a 50-unit long repeat expression vector. However, knowing that repeat expression by themselves can be highly toxic, we turned to a transient expression model. The RNA interference (RNAi) system in *C. elegans* is highly characterized and *C. elegans* is one of the only organisms that can easily uptake and amplify the RNA. Since the GGGGCC repeat is not conserved in the *C. elegans* orthologue of *C9orf72*, *alfa-1*, the expression of a GGGGCC RNA would not cause the degradation of a target RNA, but rather result in the expression of the RNA strand in most cells of the organism. BLAST searches did not identify any GGGGCC$_2$ sequence in the *C. elegans* genome and the CCCCGG$_2$ sequence (the anti-sense transcript) only blasted to a non-coding region outside of *hpo-25* (clone F16G10.4), which suggest that expression of the GGGGCC$_{50}$ sequence should not target any gene in *C. elegans* for degradation. The repeat was cloned into the L4440 vector, which contains two T7 promoters allowing for the production of the RNA encoding the GGGGCC repeat containing 50 units, but not of the protein encoded (Supplementary Fig. 1 A).

In *Drosophila*, the expression of a pathogenic GGGGCC repeat was shown to affect neuronal integrity [7-9]. Locomotion of *C. elegans* was often shown to provide a good

evaluation of neuronal integrity [10-12]. Therefore, the toxicity of the GGGGCC repeat was evaluated by looking at its effect on worm's motility. First, RNAi resistant strains were tested. These mutant strains have mutation in major component of the RNAi processing machinery make them resistant to RNAi, so there should not be a difference between the toxicity of the $GGGGCC_{50}$ and the one caused by the GFP RNA. Both strains were resistant to the $GGGGCC_{50}$ RNA, and worms did not exhibit paralysis rates higher than the GFP RNA used as control, even at day 12 of adulthood (Supplementary Fig. 1 B, C and Supplementary Table 1A-B). We have previously demonstrated that GABAergic neurons were shown to be sensitive to an *alfa-1* loss of function toxicity[13], so mutant animals genetically sensitized to RNAi-effect only the GABAergic neurons were used. Exposure of the GABAergic neurons to $GGGGCC_{50}$ RNA caused age-dependant motility problems leading to paralysis of the animals (Fig. 1 A and Supplementary Table 1C). Also, using a *unc-47p*::mCherry marker to visualize the GABAergic neurons, breaks along the axons were observed at days 1, 5 and 9 and increased with the age of the animals (Fig. 1 B, C), confirming that exposure to $GGGGCC_{50}$ RNA affects neuronal integrity in an age-dependent manner.

Finally, to compare our model with previous *C9orf72* models characterized, animals were submitted to TMPγP4 (5,10,15,20-Tetrakis(1-methyl-4-pyridinio)porphyrin tetra(p-toluenesulfonate)). TMPγP4 binds to the G-quadruplex RNA structure and was shown to alleviate the toxicity of the expanded GGGGCC repeat in model organisms [9,14]. In *C. elegans*, TMPγP4 was demonstrated to alleviate the motility impairment and neurodegeneration caused by the exposure to $GGGGCC_{50}$ at different concentrations (Fig. 1 D, E and Supplementary Table 1D), confirming that our model recapitulates key features of *C9orf72*

GGGGCC repeat expansion toxicity. Taken together, our data indicate that exposure to $GGGGCC_{50}$ RNA is toxic to *C. elegans* neurons and could be alleviated with drugs acting on GGGGCC repeat toxicity.

$GGGGCC_{50}$ RNA and *alfa-1* loss of function are toxic to different tissues

Using a deletion mutant worm strain, we previously showed that decreased expression of the *C9orf72* orthologue, *alfa-1*, was detrimental to *C. elegans* movement. Using newly available transgenic strains, we reevaluated the effect of a decreased expression of *alfa-1* in different tissues using RNAi. Interestingly, *alfa-1* RNAi caused motility problems in *C. elegans* when *alfa-1* expression was decreased in neuronal, muscle and intestinal tissues (Supplementary Fig. 2 A, B, C), suggesting that a decreased expression of *alfa-1* is detrimental to many types of tissue in *C. elegans*.

To determine the toxicity of the $GGGGCC_{50}$ RNA, we used the same RNAi-sensitive strains to monitor the effect of the transport of $GGGGCC_{50}$ RNA into different cell types of the animal. Exposure to pathogenic $GGGGCC_{50}$ RNA did not cause paralysis when using animals sensitized to RNAi-effects in muscle or intestinal cells (Supplementary Fig. 2 F, E), but caused high level of paralysis to animals sensitized to RNAi effects in the nervous system (Supplementary Fig. 2 D). When evaluated as a ratio of worms paralyzed at day 12 submitted to target RNA/control RNA, our data suggest that decreased expression of *alfa-1* resulted in paralysis (ratio >1) in all tissues, while the susceptibility to $GGGGCC_{50}$ RNA only caused paralysis in neuronally-sensitized animals, while the RNAi resistant strains showed no overt paralysis when exposed to $GGGGCC_{50}$ RNA (Fig. 2). These data indicate that in

C. elegans, the $GGGGCC_{50}$ RNA is toxic only to neurons while a decreased in *alfa-1* is detrimental to many types of cells.

Development of a rapid high throughput drug screening

C. elegans is usually cultured on Petri plates where they crawl across the surface as the main form of motility. However, worms can be grown in liquid culture where they exhibit a stereotypical swimming motion. Using our different ALS models, we previously showed that when the worms are put in liquid culture, the paralysis that is observed after days when they are grown on plates, is highly accelerated and can be detected after only few hours in liquid culture [13,15]. The swimming activity of the animal actively engages the neuromuscular junction and may be a good way to evaluate motor neuron health. We have previously shown that the *alfa-1(ok3062)* animals exhibit specific neurodegeneration affecting the motor neurons and abnormal neuromuscular junction function[13]. Therefore we sought to evaluate their movement in liquid culture. Worms were placed in a 96-well plate and their movements were quantified with an automated method that measures locomotion activity based on infrared beam scattering [16] . After only 30 minutes in liquid culture, a difference between the wild-type N2 worms and *alfa-1(ok3062)* was observed, where the *alfa-1(ok3062)* mutants exhibit less movement. This phenotype is maintained during up to six hours and is suitable for phenotypic screening (Fig. 3 A).

Using this automated method we have developed a drug-screening platform. Young adults were put in liquid culture with control (DMSO) or with compounds (at 20 µM) and their movement was evaluated for two hours. More than 4,000 bioactive molecules from

several commercialised drug libraries were screened. After two hours in liquid culture, 80 compounds were shown to affect significantly motility of the *alfa-1 (ok3062)* worms (Supplementary Table 3). From these 80 drugs, 25% are known to act on the central nervous system or at the neuromuscular junction in humans and include mainly antipsychotic drugs as well as antidepressants, antispastics, anticonvulsants, analgesics, antiemetics and central nervous system stimulants (Fig. 3 B)

A similar screen was previously done in our laboratory using the same method to identify compounds that could alleviate the toxicity observed in transgenic *C. elegans* models expressing TDP-43 or FUS ALS causing mutations or by the deletion of *kcc-2* in *C. elegans* (unpublished results). The transgenic TDP-43^{A315T} and FUS$^{S57\Delta}$ worms also exhibit age-dependent paralysis and neurodegeneration of the GABAergic motor neuron probably due to high proteotoxic burden[11]. Few similarities are observed between the different drug-screening data sets. *alfa-1* deletion mutation and TDP-43 transgenic worms share only three common target drugs, which included pizotifen maleate, cyproheptadine hydrochloride and melatonin (Fig. 3 C). The *kcc-2* deletion mutant exhibits motility impairments due to GABAergic neuron developmental problems (unpublished data and [17]). Also, the movement impairments of *alfa-1* and *kcc-2* mutants are rescued by only one common compound, chlorpromazine (Fig. 3 C). These data confirmed that the drugs identified to increase the motility of *alfa-1* deletion mutants are highly specific to *alfa-1* toxicity.

Small molecules rescue neuromuscular hypersensitivity and neurodegeneration in *alfa-1(ok3062)* mutants

Knowing that the liquid environment in which the screen was performed is not the normal conditions in which *C. elegans* are assessed, 11 of the 80 drugs were chosen for characterization on solid media. The 11 drugs were chosen based on their efficacy and included melatonin and pizotifen maleate. Melatonin and pizotifen maleate also were both shown to alleviate transgenic TDP-43^{A315T} toxicity (data not shown) and melatonin was previously demonstrated to alleviate psychiatric and age-associated disorders[18,19] . On solid media, we have previously shown that the *alfa-1(ok3062)* animals exhibit age-dependent paralysis, and neurodegeneration of the GABAergic motor neurons[10]. Also, young adult animals exhibit hyperexcitability to aldicarb, a compound used to evaluate the function of the neuromuscular junction [10].

All molecules were tested at 20 μM, with the exception of zuclopenthixol dihydrochloride which was toxic at this concentration and was therefore used at 2 μM. On solid plates, all 11 compounds were shown to significantly reduce the number of worms that were paralyzed on day 12 of adulthood (Fig. 3D). Furthermore, all of them rescued the neurodegeneration observed at day 9 in the *alfa-1(ok3062)* animals (Fig. 3 D, E). Finally, the majority of the compounds (9/11) rescued the hypersensitivity to aldicarb observed in the *alfa-1(ok3062)* animals (Figure 3 F). These data therefore confirmed that the screen done in liquid culture could identify compounds that affect *C. elegans* age-dependent loss of motility and neurodegeneration.

Small molecules rescue of $GGGGCC_{50}$ RNA toxicity

In order to determine if the molecules identified in previous screen could also affect the toxicity caused by the exposure of the $GGGGCC_{50}$ RNA, all eleven drugs were tested in the animals sensitive to RNAi in GABAergic neurons and exposed to $GGGGCC_{50}$ RNA. Only eight of these molecules rescued the paralysis phenotype including: prednicarbate, perphenazine, thioridazine hydrochloride, tridihexethyl chloride, milirinone, tremorine dihydrochloride, pizotifen malate and melatonin (Fig. 4A and Supplementary Fig. 3). However, only two compounds had a protective effect against neurodegeneration (Fig. 4 B). None of these compounds had an effect on the growth of the bacteria that deliver the $GGGCC_{50}$ RNA by itself (data not shown).

Interestingly, four of these compounds, perphenazine, pizotigen malate, thioridazine hydrochloride and tremorine dihydrocloride, also had an influence on the motility of wild type N2 worms (Fig. 4 C), suggesting that they act to generally maintain neuronal integrity and neuromuscular junction function.

Abnormal RNA expression in cells derived from patients carrying *C9orf72* expansions

RNA dysregulation is speculated to be a key aspect of ALS pathogenesis [20]. Additionally, the formation of RNA foci and production of dipeptides caused by the presence of the expanded GGGGCC RNA were also reported to lead to abnormal RNA expression [21-24] . Therefore, we sought to characterize the transcriptome of cell lines derived from ALS patients presenting *C9orf72* expansions. In order to do so, whole-transcriptome sequencing (RNA-Seq) was prepared from *C9orf72* positive fibroblasts, along wild-type control

fibroblasts. Additional RNA-Seq was carried also from lymphoblastoid cells derived from the very same individuals from which the fibroblasts were derived. Such an approach was used to identify changes in expression that would be due to the presence of expanded GGGGCC *C9orf72* and not to a cell specific expression profile. Additionally, we further examined RNA expression in two additional *C9orf72* positive fibroblast cell lines for which lymphoblastoid cell lines were not available. For each sample, a minimum of 50 million reads mapped to the human genome. An analysis of the different data sets revealed that 130 transcripts had a differential expression with a non-adjusted p value < 0.05 (Supplementary Table 4). Functional annotation using DAVID annotation tool identified many clusters of dysregulated transcripts including signal peptide, glycoprotein and splice variant which is similar to the observed enrichment identified in sALS [25]

Molecules identified restored abnormal gene expression

We next evaluated the effect of the molecules identified in *C. elegans* for their activity using ALS patient derived cell lines. Three transcripts for which expression was dysregulated in *C9orf72* positive cells were selected. These transcripts are: *beta-secreatse-2* (*BACE2),* which is involved in the processing of amyloid precursor protein, *interleukin-6 (IL6)*, which was shown to be involved in the immune system in ALS, and *protocadherin-gamma A4*(*PCDHGA4),* which plays a role in RNA processing[26,27]. These transcripts were chosen because they are expressed at high levels in fibroblast cell lines and their functions could be relevant to ALS pathogenesis. All three transcripts were observed to be upregulated in cells from patients carrying an expanded GGGGCC repeat (Fig. 5 B).

As a proof of principle TPMγP4, a molecule that was shown to disrupt the toxic G-quadruplex structure of the expanded GGGGCC RNA, was tested to evaluate if it could restore the abnormal expression of those three transcripts. Three lines of fibroblast cell lines were submitted to TMPγP4 for 72 hours and dose-response was observed in cells treated with TMPγP4 compared to the untreated controls (Fig. 6 B). Thus, we speculated that the RNA abnormal expression could be a good phenotype to evaluate *C9orf72* toxicity.

The eight molecules that were shown to alleviate the toxicity of both the decreased expression of *alfa-1* and the susceptibility to $GGGGCC_{50}$ RNA were tested in three fibroblastoid cell lines. From these eight molecules, prednicarbate and tridihexethyl chloride were found to restore the normal expression of *BACE2*, *IL6* and *PCDHGA4* (Fig. 6 C, D, Supplementary Fig 4). Our results demonstrate that some of the drugs identified in *C. elegans* also alleviate toxic RNA metabolism observed in *C9orf72* positive patient cells.

DISCUSSION

C9orf72 is the most common known cause of ALS and many toxic mechanisms were speculated regarding its toxicity. A decreased expression of *C9orf72*, the formation of nuclear RNA foci and the expression of many dipeptides were all observed in patient neuronal cells[28-31]. To understand more about these modes of toxicity, model organisms are essential. Using the worm *C. elegans*, we developed models to understand the toxicity of a decreased expression, as well as the toxicity caused by the GGGGCC repeat RNA. Our data suggest that both means of toxicity are toxic to *C. elegans* neurons. Because of its small size, its short generation time and the high conservation of many genetic and cellular pathways

with humans, *C. elegans* is an increasingly popular choice in organism-based high throughput screening [32]. Therefore, having different worms that recapitulated *C9orf72* toxicity, we carried out a drug screening to identify molecules that could alleviate both types of toxicity. Eight drugs were shown to attenuate the motility impairments observed in both models and two molecules were also shown to restore the abnormal RNA expression observed in patient cells. We speculated that these drugs could be promising for ALS patients and should be further tested in other animal models of *C9orf72*. Finally, the characterization of their effect could highlight pathogenic mechanisms regarding *C9orf72* and its expanded GGGGCC repeat.

From the eight molecules that could alleviate motor impairment in both *C. elegans* models, pizotifen malate and melatonin were also shown to mediate the motor impairments caused by the expression of mutant TDP-43 and FUS in *C. elegans*. Pizotifen malate was also shown to increase the movement of the N2, wild-type worms, suggesting that it may act in a non-specific manner to maintain neuromuscular junction health. It was previously shown that drugs acting specifically targeting the function of the neuromuscular junction could be promising to treat neuromuscular disorders[33].

Of interest, melatonin was also observed to modulate phenotype in animal models of other neurodegenerative disorders including Alzheimer's disease and Parkinson's disease [34,35]. Melatonin is synthesized from tryptophan and is known to participate to many cellular processes including immune function, cell growth, circadian rhythms and free radical scavenging[19]. Melatonin's protective role in neurodegenerative disorders is speculated to be due to its activity as an antioxidant [19]. Since the identification of mutations

in *SOD1* in fALS cases, mitochondrial dysfunction and reactive oxygen species have been a recurrent theme in ALS pathogenesis. Melatonin was shown to attenuate the phenotypes observed in *SOD1* mouse model and was even demonstrated to be safe in ALS patients[36]. Even though melatonin was only found to affect *C9orf72* neuronal toxicity, it could be speculated that it acts specifically on neuronal cells. Interestingly, Zhang *et al.* suggested that neurotoxicity of *C9orf72* dipeptides induced endoplasmic reticulum stress (ER stress)[37] and a cross-talk between ER and mitochondria, that would affect ROS generation, is speculated in ALS [38], perhaps explaining the role of melatonin in *C9orf72* toxicity. These data suggest that melatonin could influence the neurotoxicity caused by many ALS mutant proteins and its effect in patients should be re-evaluated.

The two drugs that restored the phenotypes in *C. elegans* and in patient cells included prednicarbate and tridihexethyl chloride. Tridihexethyl chloride is an anticholinergic agent that used to be taken to treat eye movement[39]. It was reported to have many side effects so it is no longer in use. However, either it's anticholinergic effect, or other off target effects specifically affects *C9orf72* toxicity suggesting that a characterization of its interacting partners might highlight major protective pathways regarding *C9orf72* toxicity. Also, it is important to note that riluzole, the only drug that is approved to treat ALS, was also shown to be an anticholinergic agent[40]. Therefore, suggesting that the effect of anticholinergic agents might be more protective in *C9orf72* positive patients.

The pathogenic molecular mechanism of *C9orf72* is still unclear. The loss of expression, and the RNA containing the expanded repeat are both toxic in our models.

Furthermore the mechanisms of TDP-43 or FUS toxicity are not fully resolved since both a gain, and a loss of function are speculated to be toxic [6,41]. Moreover, for some repeat disorders, a gain of function at the protein and/or RNA levels is often observed, and this can be accompanied by a loss of the normal function of the protein[42]. Our data indicate that a similar situation is observed in *C9orf72* where both would be highly toxic to neurons but the loss of *C9orf72* would be the most detrimental to non-neuronal cells. The drugs identified in our screen also highlight the importance of the ER stress and mitochondrial dysfunctions regarding *C9orf72* pathogenesis and suggest that anticholinergic drugs, similar to riluzole, might be more effective in *C9orf72* positive patients.

METHODS

C. elegans strains and maintenance

Standard method for culturing and handling the worms were used[43]. Worms were cultured on standard NGM media streak with OP-50 Escherichia and maintained at 20°C if not specified otherwise. For a list of strains used see Supplementary Table 3

RNAi experiments

RNAi-treated strains were fed with E. coli(HT115) containing an empty vector (EV) or *alfa-1* (F18A1.6) RNAi clones from the ORFeome RNAi library (Open Biosystems). RNAi experiments were performed at 20°C. Worms were grown on NGM enriched with 1mM Isopropyl-b-thiogalacto-pyranoside (IPTG).

Construction of the $GGGGCC_{50}$ RNAi expression vector and worm maintenance

Generation of the $GGGGCC_{50}$ was created by annealing oligonucleotide sequence of GGGGCCGGGGCCGGGGCC at 37C. The generated longer repeat was purified on gel and inserted into a pBLueScript vector (Addgene) containing HA, His and 6x Myc in the three different potential reading frames. The $GGGGCC_{50}$ and the tags were then cloned into the L4440 vector (Addgene) that was transformed into HT115 *E. coli*. Worms were fed from hatching with the $GGGGCC_{50}$ containing bacteria and transferred every three days on fresh plates.

Paralysis assay

Worms were transferred on 5μM fluorodeoxyuridin (FUDR) plates one day after L4. Worms were scored daily for movement for 12 days. Worms were counted as paralyzed if they failed to move after they were prodded on the nose. Experiments were performed at 20°C and at least 60 worms were counted per conditions. Survival curves and statistics were produced using Log-rank (Mantel-Cox) test using GraphPad Prism software.

Liquid culture assay

A synchronized population was obtained using hypochlorite extraction. Worms were grown on solid media up to day 1 of adulthood. At day 1, 50 worms per well were placed in S basal with OP-50 E. coli (optical density 0.5) in a flat-bottom 96-well plate. Standard errors are shown on the graph. Measurement was done using Microtracker (Phylumtech) with standard parameters for *C. elegans*.

Drug screening

Preparation of the animals for liquid culture was made as described above. Drugs included the following libraries: SIGMA Lopas, Microsource Discovery Spectrum, Biomol Natural Products and the Prestwick Commercialized products libraries. Drugs were received frozen and were kept in DMSO. All drugs were first tested in 1 well/drug condition at concentration of 20uM for 2 hours. Drugs that have caused an increased in *alfa-1(ok3062)* movement were then retested in triplicate at the same concentration and time. Measurement was done using Microtracker (Phylumtech) with standard parameters for *C. elegans*.

Neuronal integrity

Animals were selected at day 1, 5 or 9 of adulthood (post L4 stage) for *in vivo* visualisation. Animals were immobilised in 5 mM levamisole and mounted on 2% agarose pads. Neurons were visualised with a Zeiss Axio Imager M2 microscope. The software used was Zen Pro 2012. At least 30 worms were counted per conditions. Statistics were produced using Log-rank (Mantel-Cox) test using GraphPad Prism software.

Cell culture

Lymphoblastoid cell lines (LCL) from ALS and controls individuals were grown in IMDM (Gibco) supplemented with foetal bovine serum (10%), fungizone (1.25 ug/ml), penicillin & streptomycin (100 units/ml) and L-glutamine (0.292 mg/ml). The fibroblasts were

derived from skin biopsies of control individuals as described by Villegas and McPhaul[21] and were grown in MEM (Gibco) supplemented with foetal bovine serum (20%), fungizone (1.25 ug/ml), penicillin & streptomycin (100 units/ml).

For cells treatment with different molecules, 0.2 μM, 2 μM and 20 μM concentrations were mixed in the cell medium for 72h and then collected for RNA extraction.

RNA extraction

RNAs were extracted using trizol/chloroform protocol (Ambion), for q-RT-PCR with Taqman probes, or RNeasy kit (Qiagen), for RNA-sequencing.

cDNA and Quantitative-PCR (Q-PCR)

For all experiments, cDNA was produced using up to 1.25 ug of RNA and 1 ul of SuperScript enzyme in a final volume of 10 ul (SuperScript Vilo cDNA synthesis, Invitrogen).

Quantitative gene expression analysis of *C9orf72* was performed using Taqman probe assays (*C9orf72* universal probe, Hs00376619_m1) and Taqman 2X universal PCR master mix (Life Technologies). The fluorescence was read with the 7900HT Fast Real-Time PCR System from Applied Biosystems. Each assay was conducted in three replicates and the delta Cycle threshold (Ct) method was used to assess the relative quantification (RQ) of

C9orf72. The endogenous gene used for the calculation of the delta Ct was *POLR2A* (*POLR2A* probe, Hs00172187_m1).

Whole-transcriptome sequencing

2ug amount of RNA was sent for RNA-sequencing. CDNA libraries and sequencing was performed at the McGill Innovation center (Montreal, Canada). cDNA libraries were prepared with the firststrand TrueSeq mRNA protocol of Illumina and sequencing was done on the Illumina Hiseq 2000/2500 sequencer. Sequencing generated between 42-173 million paired read/library.

Resultant reads were aligned to GRCh37 human genome with the STAR-rna software in a 2 pass-mode. Reads were then assemble using Cufflinks program[44] and differential gene expression analysis was done using DESeq [45] and edgeR [46] R bioconductor package

REFERENCE

1. Kiernan, M. C. *et al.* Amyotrophic lateral sclerosis. *The Lancet* **377,** 942–955 (2011).
2. Mitsumoto, H., Brooks, B. R. & Silani, V. Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? *Lancet Neurol* **13,** 1127–1138 (2014).
3. Lattante, S., Ciura, S., Rouleau, G. A. & Kabashi, E. Defining the genetic connection linking amyotrophic lateral sclerosis (ALS) with frontotemporal dementia (FTD). *Trends Genet.* **31,** 263–273 (2015).
4. Guerreiro, R., Brás, J. & Hardy, J. SnapShot: Genetics of ALS and FTD. *Cell* **160,** 798–798.e1 (2015).
5. Renton, A. E., Chiò, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci* **17,** 17–23 (2014).
6. Ling, S.-C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79,** 416–438 (2013).
7. Mizielinska, S. *et al.* C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science* 1256800 (2014). doi:10.1126/science.1256800
8. Freibaum, B. D. *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525,** 129–133 (2015).
9. Zhang, K. *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525,** 56–61 (2015).
10. Therrien, M., Rouleau, G. A., Dion, P. A. & Parker, J. A. Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in C. elegans. *PLoS ONE* **8,** e83450 (2013).
11. Vaccaro, A. *et al.* Mutant TDP-43 and FUS cause age-dependent paralysis and neurodegeneration in C. elegans. *PLoS ONE* **7,** e31321 (2012).
12. Maruyama, I. N. & Brenner, S. A phorbol ester/diacylglycerol-binding protein encoded by the unc-13 gene of Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U.S.A.* **88,** 5729–5733 (1991).
13. Therrien, M., Rouleau, G. A., Dion, P. A. & Parker, J. A. Deletion of C9ORF72 Results in Motor Neuron Degeneration and Stress Sensitivity in C. elegans. *PLoS ONE* **8,** e83450 (2013).
14. Zamiri, B., Reddy, K., Macgregor, R. B. & Pearson, C. E. TMPyP4 porphyrin distorts RNA G-quadruplex structures of the disease-associated r(GGGGCC)n repeat of the C9orf72 gene and blocks interaction of RNA-binding proteins. *Journal of Biological Chemistry* **289,** 4653–4659 (2014).
15. Vaccaro, A. *et al.* Methylene blue protects against TDP-43 and FUS neuronal toxicity in C. elegans and D. rerio. *PLoS ONE* **7,** e42117 (2012).
16. Simonetta, S. H. & Golombek, D. A. An automated tracking system for

Caenorhabditis elegans locomotor behavior and circadian studies application. *J. Neurosci. Methods* **161,** 273–280 (2007).
17. Tanis, J. E., Bellemer, A., Moresco, J. J., Forbush, B. & Koelle, M. R. The potassium chloride cotransporter KCC-2 coordinates development of inhibitory neurotransmission and synapse structure in Caenorhabditis elegans. *J. Neurosci.* **29,** 9943–9954 (2009).
18. Hardeland, R. Melatonin and Synthetic Melatoninergic Agonists in Psychiatric and Age-associated Disorders: Successful and Unsuccessful Approaches. *Curr. Pharm. Des.* (2015).
19. Pandi-Perumal, S. R. *et al.* Melatonin antioxidative defense: therapeutical implications for aging and neurodegenerative processes. *Neurotox Res* **23,** 267–300 (2013).
20. Polymenidou, M. *et al.* Misregulated RNA processing in amyotrophic lateral sclerosis. *Brain Research* **1462,** 3–15 (2012).
21. Sareen, D. *et al.* Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci Transl Med* **5,** 208ra149–208ra149 (2013).
22. Kwon, I. *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345,** 1139–1145 (2014).
23. Donnelly, C. J. *et al.* RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron* **80,** 415–428 (2013).
24. Lagier-Tourenne, C. *et al.* Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proceedings of the National Academy of Sciences* **110,** E4530–9 (2013).
25. Jones, A. R. *et al.* Stratified gene expression analysis identifies major amyotrophic lateral sclerosis genes. *Neurobiology of Aging* **36,** 2006.e1–9 (2015).
26. Berjaoui, S. *et al.* Complex Inflammation mRNA-Related Response in ALS Is Region Dependent. *Neural Plast.* **2015,** 573784–11 (2015).
27. Hussain, I. *et al.* ASP1 (BACE2) cleaves the amyloid precursor protein at the beta-secretase site. *Mol. Cell. Neurosci.* **16,** 609–619 (2000).
28. Renton, A. E. *et al.* A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **72,** 257–268 (2011).
29. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72,** 245–256 (2011).
30. Mori, K. *et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339,** 1335–1338 (2013).
31. Ash, P. E. A. *et al.* Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77,** 639–646 (2013).
32. O'Reilly, L. P., Luke, C. J., Perlmutter, D. H., Silverman, G. A. & Pak, S. C. C. elegans in high-throughput drug discovery. *Adv. Drug Deliv. Rev.* **69-70,** 247–253 (2014).
33. Arimura, S. *et al.* Neuromuscular disease. DOK7 gene therapy benefits mouse

models of diseases characterized by defects in the neuromuscular junction. *Science* **345,** 1505–1508 (2014).
34. Rudnitskaya, E. A. *et al.* Melatonin Attenuates Memory Impairment, Amyloid-β Accumulation, and Neurodegeneration in a Rat Model of Sporadic Alzheimer's Disease. *J. Alzheimers Dis.* **47,** 103–116 (2015).
35. Carriere, C. H., Kang, N. H. & Niles, L. P. Chronic low-dose melatonin treatment maintains nigrostriatal integrity in an intrastriatal rotenone model of Parkinson's disease. *Brain Research* **1633,** 115–125 (2015).
36. Weishaupt, J. H. *et al.* Reduced oxidative damage in ALS by high-dose enteral melatonin treatment. *J Pineal Res* **41,** 313–323 (2006).
37. Zhang, Y.-J. *et al.* Aggregation-prone c9FTD/ALS poly(GA) RAN-translated proteins cause neurotoxicity by inducing ER stress. *Acta Neuropathol* **128,** 505–524 (2014).
38. Manfredi, G. & Kawamata, H. Mitochondria and endoplasmic reticulum crosstalk in amyotrophic lateral sclerosis. *Neurobiol. Dis.* 1–8 (2015). doi:10.1016/j.nbd.2015.08.004
39. Leigh, R. J., Burnstine, T. H., Ruff, R. L. & Kasmer, R. J. Effect of anticholinergic agents upon acquired nystagmus: a double-blind study of trihexyphenidyl and tridihexethyl chloride. *Neurology* **41,** 1737–1741 (1991).
40. Deflorio, C. *et al.* Riluzole blocks human muscle acetylcholine receptors. *J. Physiol. (Lond.)* **590,** 2519–2528 (2012).
41. Sephton, C. F. & Yu, G. The function of RNA-binding proteins at the synapse: implications for neurodegeneration. *Cell. Mol. Life Sci.* **72,** 3621–3635 (2015).
42. Orr, H. T. & Zoghbi, H. Y. Trinucleotide Repeat Disorders. *Annu. Rev. Neurosci.* **30,** 575–621 (2007).
43. Stiernagle, T. Maintenance of C. elegans. *WormBook* (2006). doi:10.1895/wormbook.1.101.1
44. Roberts, A., Pimentel, H., Trapnell, C. & Pachter, L. Identification of novel transcripts in annotated genomes using RNA-Seq. *Bioinformatics* **27,** 2325–2329 (2011).
45. Anders, S. & Huber, W. Differential expression analysis for sequence count data. *Genome Biol* **11,** R106 (2010).
46. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26,** 139–140 (2010).

FIGURES

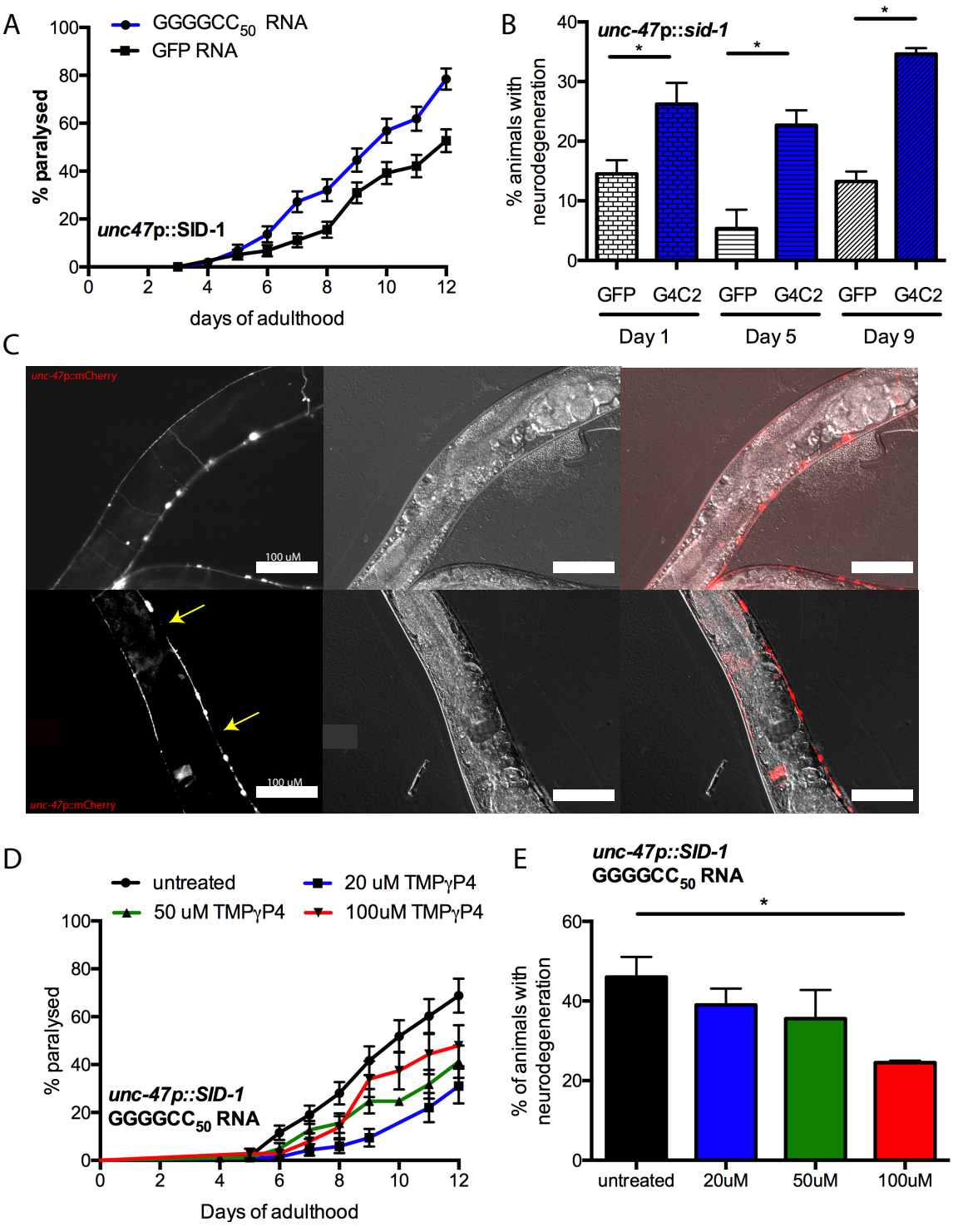

A
GGGGCC50 RNA
GFP RNA
% paralysed
unc47p::SID-1
days of adulthood
B
unc-47p::sid-1
% animals with neurodegeneration
GFP
G4C2
Day 1
Day 5
Day 9
C
unc-47p::mCherry
100 uM
D
untreated
20 uM TMPγP4
50 uM TMPγP4
100uM TMPγP4
% paralysed
unc-47p::SID-1
GGGGCC50 RNA
Days of adulthood
E
unc-47p::SID-1
GGGGCC50 RNA
% of animals with neurodegeneration
untreated
20uM
50uM
100uM

Figure 3.1: Expression of $GGGGCC_{50}$ is detrimental to *C. elegans* neurons. A) Expression of GGGGCC caused a age-dependent paralysis (A) and neuronal breaks (B) when expressed in GABEergic of *C. elegans* compared to GFP control. (C) Fluorescence, DIC and merge pictures representing the normal GABAergic neurons (upper) and an example of gaps (indicated by arrow) observed in animals exposed to GGGGCC RNA. TMPγP4 can rescue motility impairments (D) and neurodegeneration (E) caused by the expression of $GGGGCC_{50}$ RNA. * p value≤ 0.05

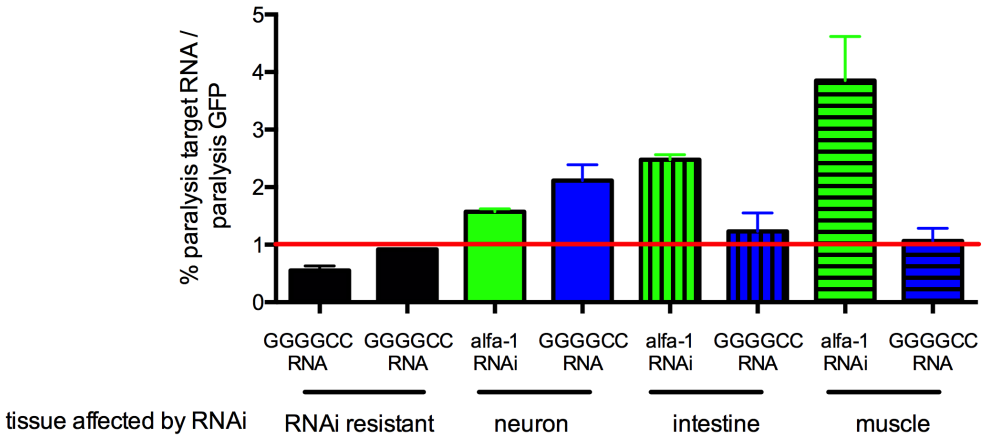
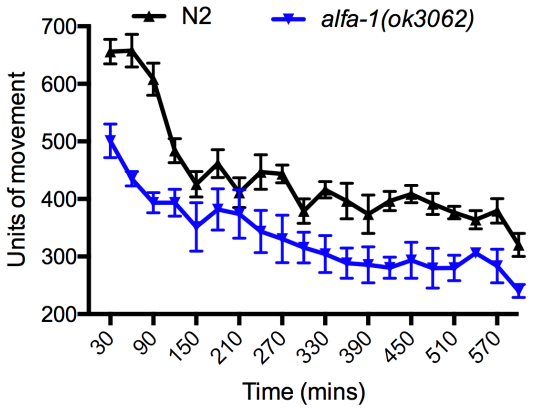


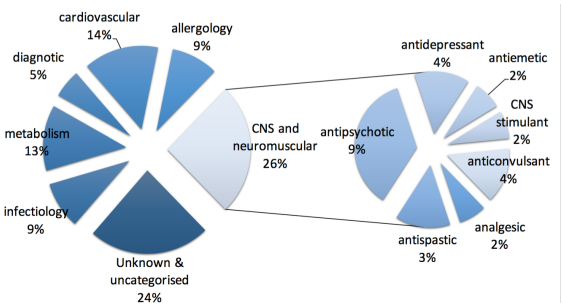
Figure 2

Figure 3.2: *alfa-1* loss of function and exposure to GGGGCC RNA are both neurotoxic. Ratio of worms paralysed at day 12 when submitted to *alfa-1* RNAi or when submitted to GGGGCC RNA in different tissues.

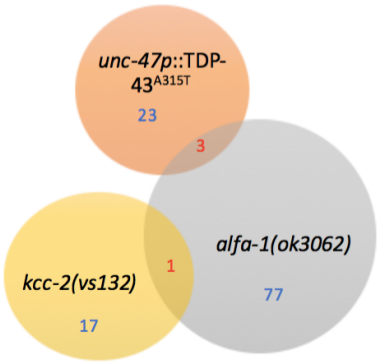
A



B



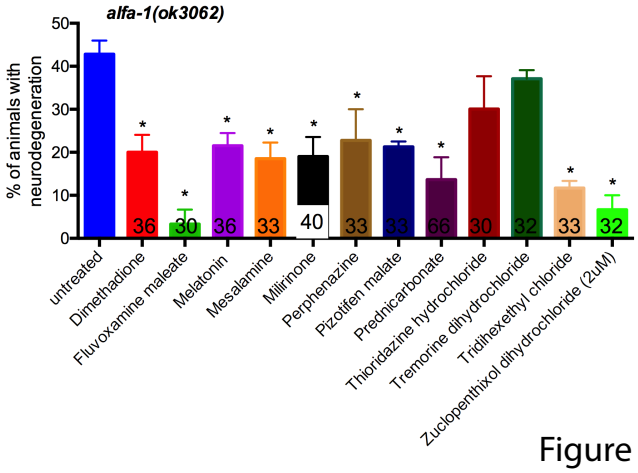
C



D

Paralysis assay	
***alfa-1(ok3062)* treatment**	Curve comparaison with untreated, p value
Dimethadione	0.0002
Fluvoxamine maleate	0.0026
Melatonin	≤0.0001
Mesalamine	0.0101
Milirinone	0.0146
Perphenazine	0.1141
Pizotifen malate	≤0.0001
Prednicarbate	0.0249
Thioridazine hydrochloride	≤0.0001
Tremorine dihydrochloride	0.0219
Tridihexethyl chloride	0.7221
Zuclopenthixol dihydrochloride 2µM	0.0159

E



F

Aldicarb assay	
***alfa-1(ok3062)* treatment**	Curve comparaison with untreated, p value
Molecules that rescue aldicarb	
Melatonin	0.0115
Mesalamine	≤0.0001
Milirinone	≤0.0001
Perphenazine	0.0022
Pizotifen malate	0.0025
Prednicarbate	≤0.0001
Thioridazine dihydrochloride	≤0.0001
Tridihexethyl chloride	≤0.0001
Molecules that do not rescue aldicarb test	
Dimetadione	0.4302
Fluzoxamine maleate	0.5583
Tremorine dihydrochloride	0.1693
Zuclopenthixol dihydrochloride 2µM	≤0.0001

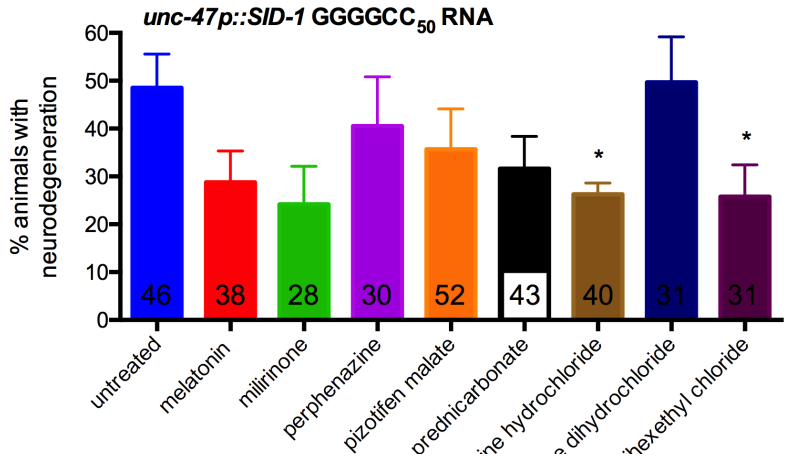
Figure 3

Figure 3.3: Molecules that alleviate *alfa-1/ C9orf72* loss of expression phenotype. (A) Liquid culture accelerates the paralysis of *alfa-1(ok3062)* worms. (B) Effect of the 80 drugs that alleviate *alfa-1(ok3062)* motor phenotype. (C) Molecules that alleviate *alfa-1(ok3062)* toxicity share few common targets with TDP-43 and FUS, and *kcc-2* mutant animals. (D) Molecules that alleviate motility impairment of *alfa-1(ok3062)* in liquid culture also alleviate the age-dependent motility phenotype on solid media. (E) Most molecules also alleviate neurodegeneration phenotype observed in the *alfa-1(ok3062)* animals. (F) Some molecules also alleviate neuromuscular junction hypersensitivity of *alfa-1(ok3062).* * p value≤ 0.05

A

unc-47p ::SID-1 GGGGCC$_{50}$ RNA Treatment	Curve comparaison with untreated, p value
Molecules that rescue GGGGCC$_{50}$ toxicity	
Melatonin	0.0123
Milirinone	≤0.0001
Perphenazine	0.0012
Pizotifen malate	≤0.0001
Prednicarbate	≤0.0001
Thioridazine hydrochloride	0.0021
Tremorine dihydrochloride	≤0.0001
Tridihexethyl chlroide	0.0284
Molecules that do not rescue GGGGCC$_{50}$ toxicity	
Dimetadione	0.1744
Mesalamine	0.4863
Fluvoxamine maleate	0.3343
Zuclopenthixol dihydrochloride	0.1418

B



C

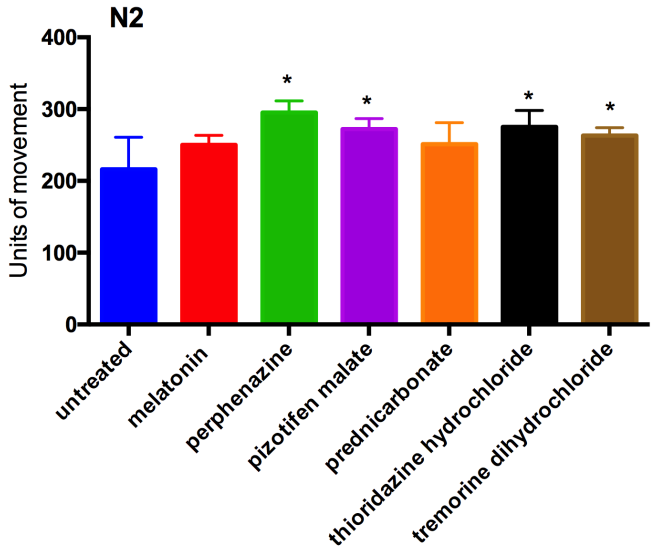


Figure 3.4: Molecules that alleviate GGGGCC$_{50}$ phenotypes. (A) Eight molecules alleviate the paralysis observed in worms exposed to GGGGCC RNA**.** (B)Thioridazine hydrochloride and tridihexethyl chloride attenuate the neuronal breaks obserbed in the GGGGCC$_{50}$ animals. (C) Perphenazine, pizotifen malate, thioridazine hydrochloride and tremorine dihydrochloride also improve motility of N2, wild-type animals. * p value≤ 0.05

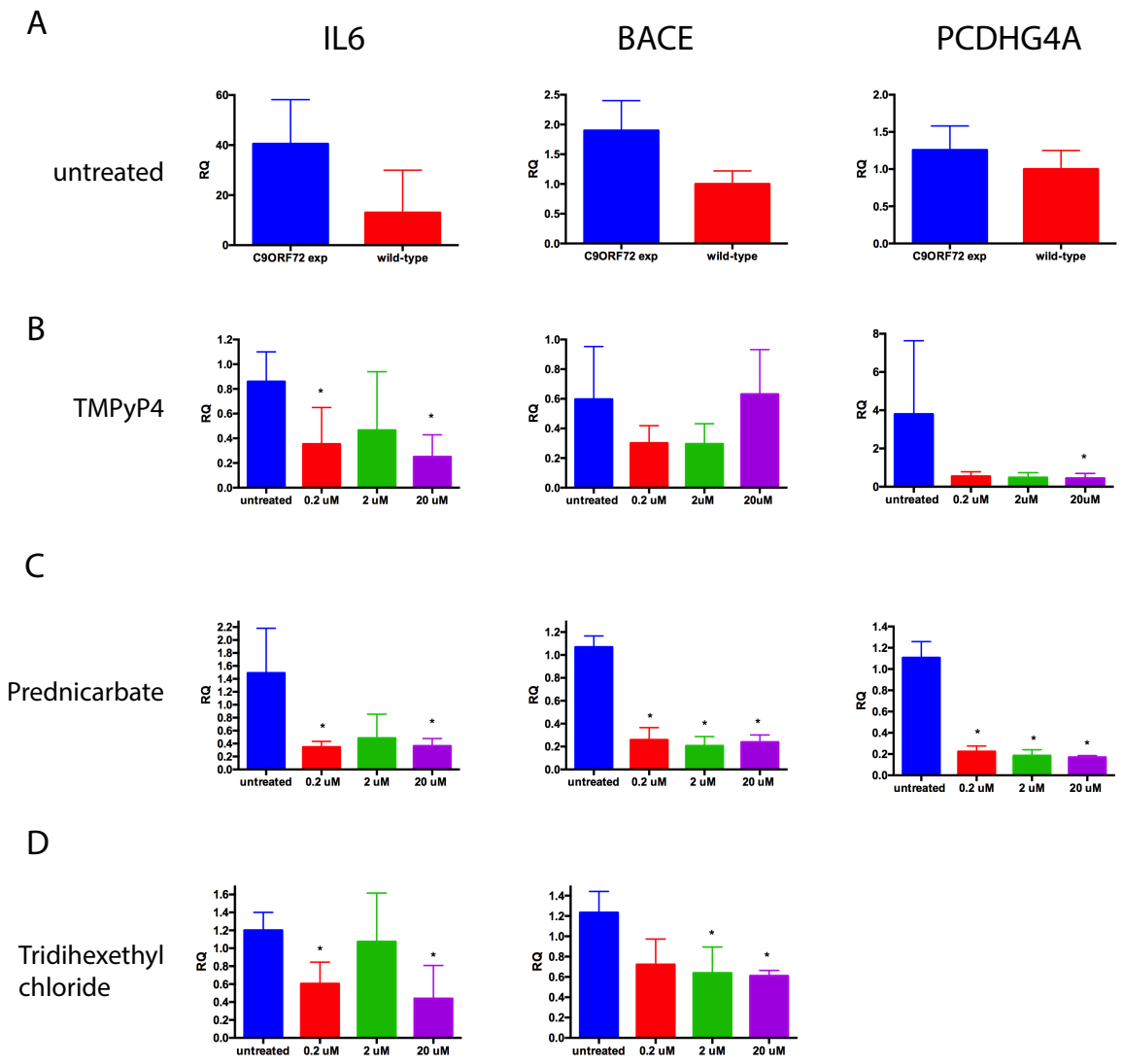
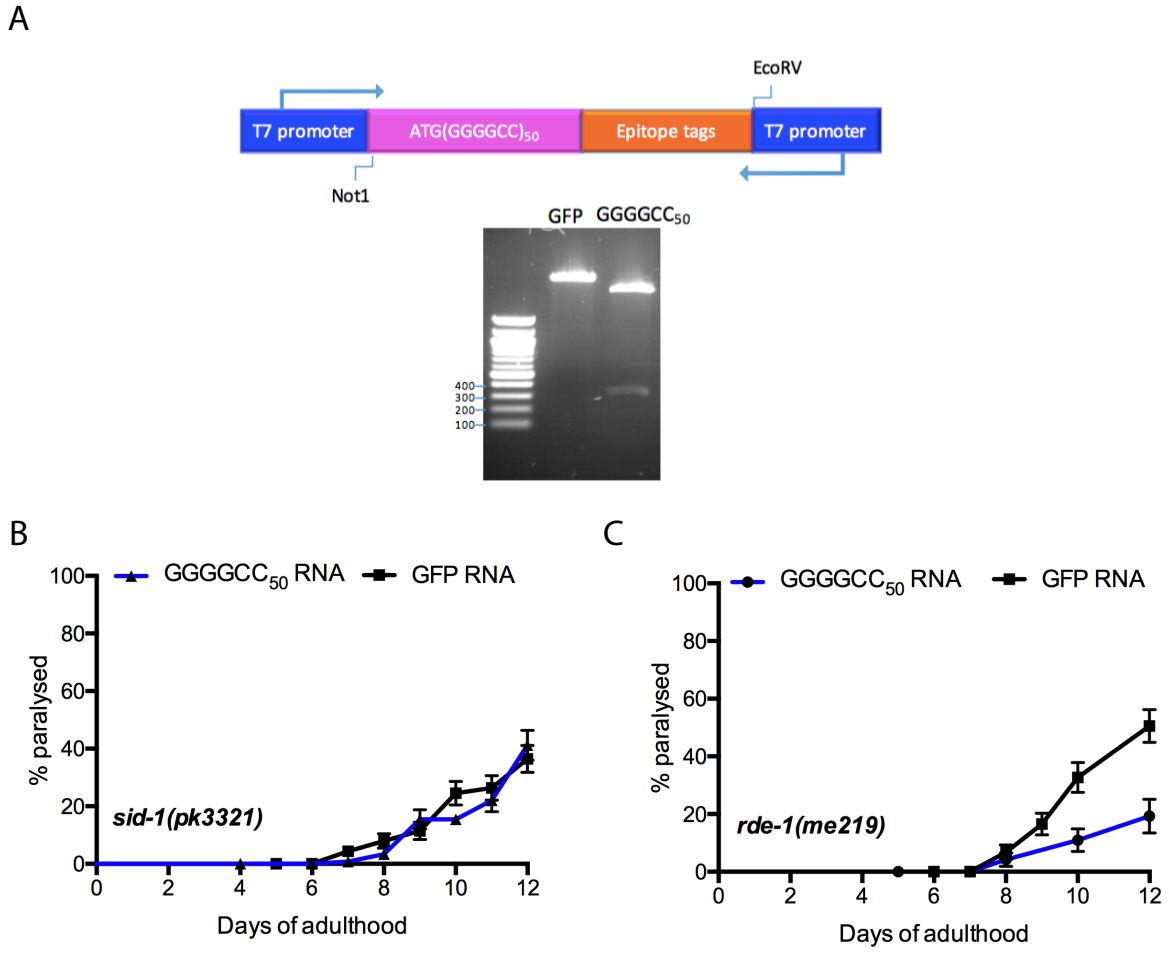
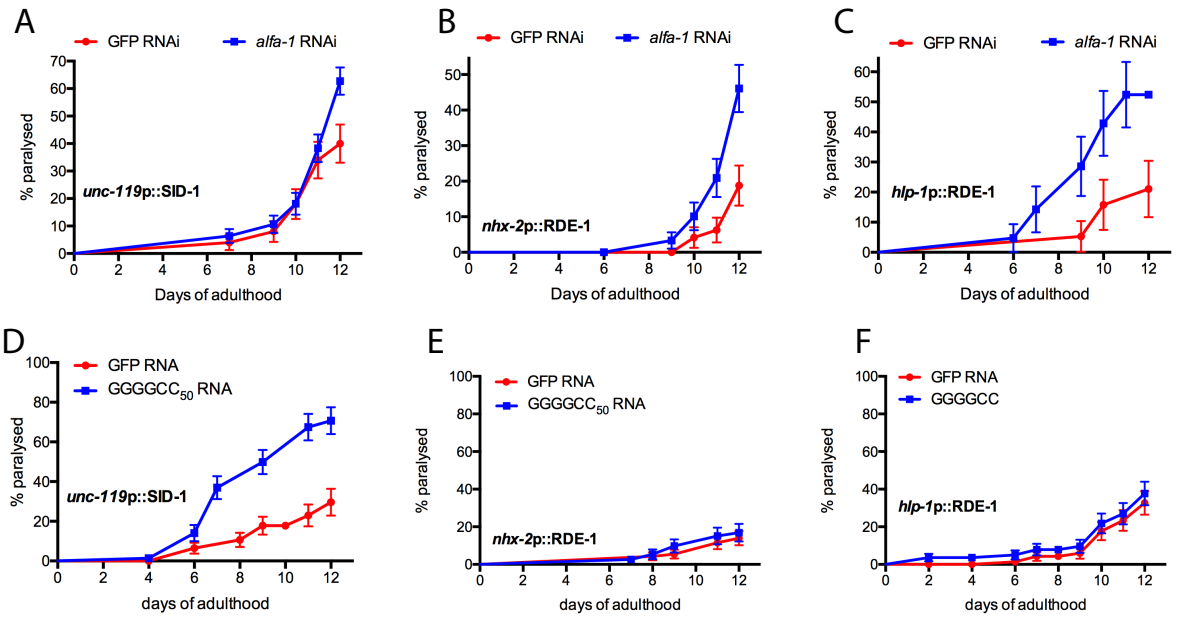


Figure 5

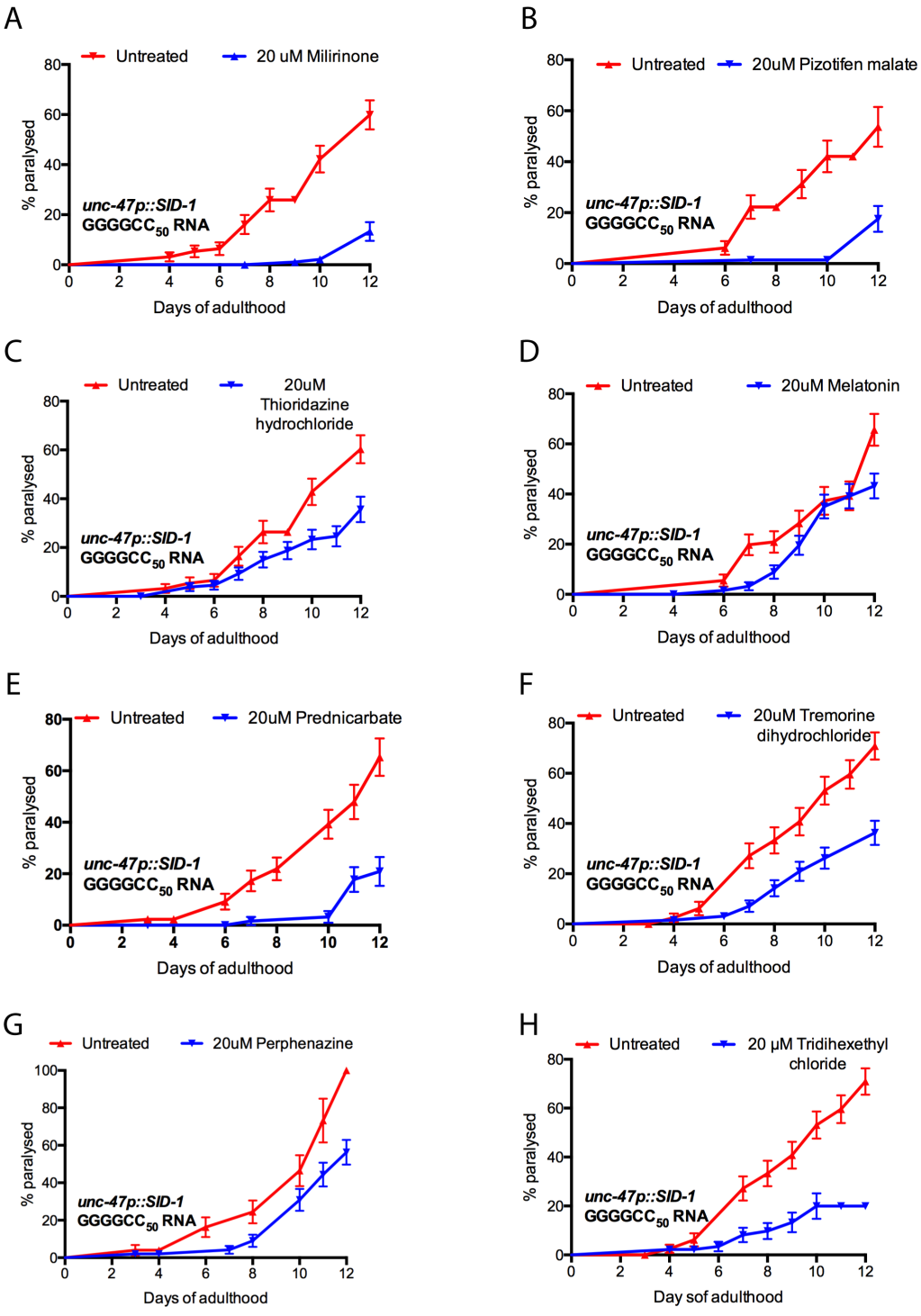
Figure 3.5: Small-molecule restoration of *C9orf72* abnormal transcriptome phenotypes (A) IL6, BACE and PCDHG4A are upregulated in *C9orf72* positive fibroblastoid cells (3 different cell lines used). (B) TMPγP4 restore that abnormal expression of IL6, BACE and PCDHG4A. (C-D). Prednicarbate and tridihexethyl chloride restore that abnormal expression of IL6, BACE and PCDHG4A. * p value≤ 0.05



Supplementary Fig. 3.1- $GGGGCC_{50}$ constructs. (A) 50 units of the GGGGCC repeat was cloned in the L4440 vector causing the expression of sense ans anti-sense transcript. Size of the repeat was confirmed by enzymatic digestion. (B-C)Exposure to GGGGCC RNA do not cause motility impairement in RNAi resistant strains

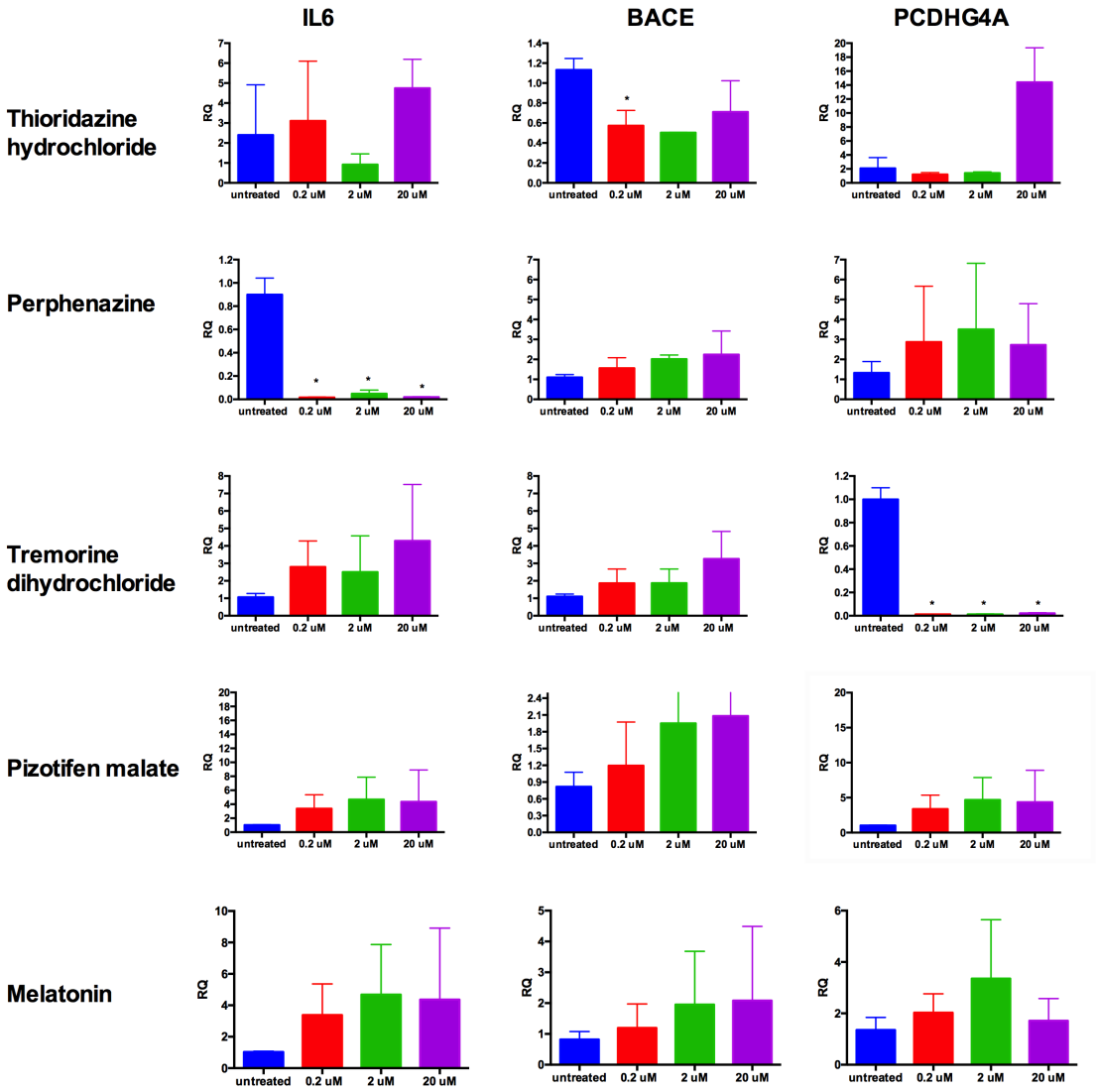


Supplementary Fig. 3.2- Tissue specific exposure to *alfa-1* RNAi and GGGGCC RNA. (A-C) *alfa-1* RNAi caused an increase paralysis when either the neuronal, muscle or intestinal cells are sensitized to RNAi effect. (D-F) exposure to GGGGCC RNA is only toxic in neuronal sensitized animals and has no effect in intestinal or muscle tissue sensitized animals.



Supplementary Figure 3

Supplementary Fig. 3.3- Molecules that can alleviate the toxicity caused by GGGGCC RNA exposure



Supplementary Fig. 3.4- Molecules that do not restore the abnormal expression of IL6, BACE and PCDHG4A

Supplementary Table 3.1

A)Statistics table for paralysis test Supplementary Fig. 1B

strain	% worms paralysed at day 12	P value
rde-1(me219) GFP RNA	50.5%	-----
rde-1(me219) $GGGGCC_{50}$ RNA	19.2%	0.0004

B)Statistics table for paralysis test Supplementary Fig. 1B

strain	% worms paralysed at day 12	P value
sid-1(pk3321) GFP RNA	36.4%	-----
sid-1(pk3321) $GGGGCC_{50}$ RNA	41.2%	0.8688

C)Statistic table paralysis Fig 1A

Strain : unc-47p ::SID-1	% worms paralysed at day 12	P value
GFP RNA	52.7%	-----
$GGGGCC_{50}$ RNA	78.4%	≤0.0001

D) Statistics for paralysis Figure 1D

treatment	% worms paralysed at day 12	P value
Untreated	68.8%	-----
20uM TMPyP4	31%	≤0.0001
50uM TMPyP4	41.1%	0.0024
100 uM TMPyP4	47.8%	0.0290

Supplementary Table 3.2

A)Statistics table for paralysis Supplementary Fig. 2A

Strain : unc-119p ::SID-1	% worms paralysed at day 12	P value
GFP RNA	40.0%	-----
alfa-1 RNAi	62.7%	0.0378

B) Statistics table for paralysis Supplementary Fig. 2B

Strain : nhx-2p ::RDE-1	% worms paralysed at day 12	P value
GFP RNA	18.7%	-----
alfa-1 RNAi	46.07%	0.0029

C)Statistics table for paralysis Supplementary Fig 2C

Strain : hlp-1p ::RDE-1	% worms paralysed at day 12	P value
GFP RNA	21.0%	-----
alfa-1 RNAi	52.38%	0.0316

D)Statistic table for paralysis Supplementary Fig 2D

Strain : unc-119p ::SID-1	% worms paralysed at day 12	P value
GFP RNA	29.6%	-----
$GGGGCC_{50}$ RNA	70.6%	≤0.0001

E)Statistic table for paralysis Supplementary Fig. 2E

Strain : nhx-2p ::RDE-1	% worms paralysed at day 12	P value
GFP RNA	13.9%	-----
$GGGGCC_{50}$ RNA	16.8%	0.5457

F)Statistic table for paralysis Supplementary Fig. 2F

Strain : hlp-1p ::RDE-1	% worms paralysed at day 12	P value
GFP RNA	32.6%	-----
$GGGGCC_{50}$ RNA	37.6%	0.5499

Supplementary Table 3.3- Molecules identified to increase *alfa-1(ok3063)* motility in liquid culture

Molecules identified to increase *alfa-1(ok3063)* motility in liquid culture			
1,2-dimethylhydrazine hydrochloride	2'-methoxyformonetin	3beta-hydroxy-23,24-bisnorchol-5-enic acid	3-hydroxyflavone
5alpha-cholestanol	8-hydroxycarapanic acid	Aminopurine, 6-benzul	Amoxicillin
Apramycin	Astemizole	Azaperone	Benzamil hydrochloride
Bromperidol	Camptothecine	Capsanthin	Cefepime hydrochloride
Chlorpromazine hydrochloride	Chlorzoxazone	Clonixin lysinate	Cloxacillin sodium salt
Colchiceine	Cyproheptadine hydrochloride	Diallyl sulfide	Diflunisal
Dimaprit dihydrochloride	Dimethadione	Dosulepin hydrochloride	Estropipate
Ethosuximide	Fenbendazole	Fendiline hydrochloride	Fluspirilien
Fluticasone propionate	Fluvastatin sodium salt	Fluvoxamine maleate	Homatropine hydrobromide (R,S)
Hydralazine hydrochloride	Indapamide	Imidurea	Iproniazide phosphate
Iopanoic acid	Isocarboxazid	Lapachol	Lomefloxacin hydrochloride
Mebendazole	Melatonin	Memantine hydrochloride	Merogedunin
Mesalamine	Methapyrilene hydrochloride	Mevastatin	Milrinone
Minocycline hydrochloride	Moricizine hydrochloride	N- (9-fluorenylmethoxycarbonyl)-L-leucine	Ononetin
Oxymetazoline hydrochloride	Ozagrel hydrochloride	Pelletierine hydrochloride	Perphenazine
Phenelzine sulfate	Pinacidil	Piperacetazine	Piperidolate hydrochloride
Pizotifen malate	Prednicarbate	Promethazine hydrochloride	Pronethalol hydrochloride
Purpurin	Quinethazone	Sertraline	Spiramycin
Streptomycin sulfate	Thioridazine hydrochloride	Tremorine dihydrochloride	Trimedlure
Trimethoprim	Trimeprazine tartrate	Tridihexethyl chloride	Ziprasidone Hydrochloride
Zuclopenthixol dihydrochloride			

Supplementary Table 3.4- Transcripts misregulated in $C9orf72$ cells : Transcripts that were misregulated in $C9orf72$ positive lymphoblast and fibroblastoid cell lines

ENSG00000254838	GVINP1
ENSG00000242588	RP11-274B21.1
ENSG00000081059	TCF7
ENSG00000175265	GOLGA8A
ENSG00000010310	GIPR
ENSG00000244479	OR2A1-AS1
ENSG00000182319	SGK223
ENSG00000204531	POU5F1
ENSG00000101605	MYOM1
ENSG00000095637	SORBS1
ENSG00000213949	ITGA1
ENSG00000254122	PCDHGB7
ENSG00000185551	NR2F2
ENSG00000178631	ACTG1P1
ENSG00000176485	PLA2G16
ENSG00000206417	H1FX-AS1
ENSG00000148798	INA
ENSG00000181031	RPH3AL
ENSG00000198208	RPS6KL1
ENSG00000236453	AC003092.1
ENSG00000109861	CTSC
ENSG00000118257	NRP2
ENSG00000203499	FAM83H-AS1
ENSG00000251580	RP11-539L10.3
ENSG00000185864	NPIPB4
ENSG00000183421	RIPK4
ENSG00000185291	IL3RA
ENSG00000133048	CHI3L1
ENSG00000182472	CAPN12
ENSG00000166689	PLEKHA7
ENSG00000165171	WBSCR27
ENSG00000162692	VCAM1
ENSG00000203867	RBM20
ENSG00000064692	SNCAIP
ENSG00000188064	WNT7B
ENSG00000105339	DENND3
ENSG00000161243	FBXO27
ENSG00000140450	ARRDC4
ENSG00000159733	ZFYVE28
ENSG00000175899	A2M
ENSG00000151079	KCNA6
ENSG00000198915	RASGEF1A
ENSG00000133424	LARGE
ENSG00000111052	LIN7A
ENSG00000243797	CTB-111H14.1
ENSG00000235505	RP11-693N9.2
ENSG00000145934	TENM2
ENSG00000062524	LTK
ENSG00000132465	IGJ
ENSG00000182022	CHST15
ENSG00000182397	DNM1P46
ENSG00000255320	RP11-755F10.1
ENSG00000215252	GOLGA8B
ENSG00000006210	CX3CL1
ENSG00000125378	BMP4
ENSG00000154864	PIEZO2
ENSG00000184232	OAF
ENSG00000109943	CRTAM
ENSG00000218537	AP000350.4
ENSG00000158089	GALNT14
ENSG00000086717	PPEF1
ENSG00000116991	SIPA1L2
ENSG00000164929	BAALC
ENSG00000131773	KHDRBS3
ENSG00000142949	PTPRF

ENSG00000141314	RHBDL3
ENSG00000136244	IL6
ENSG00000228463	AP006222.2
ENSG00000120658	ENOX1
ENSG00000157064	NMNAT2
ENSG00000088836	SLC4A11
ENSG00000135074	ADAM19
ENSG00000123243	ITIH5
ENSG00000169129	AFAP1L2
ENSG00000182621	PLCB1
ENSG00000162551	ALPL
ENSG00000105825	TFPI2
ENSG00000232745	ANKRD20A14P
ENSG00000123612	ACVR1C
ENSG00000145358	DDIT4L
ENSG00000273295	AP000350.5
ENSG00000183570	PCBP3
ENSG00000039560	RAI14
ENSG00000177875	C12orf68
ENSG00000064886	CHI3L2
ENSG00000077943	ITGA8
ENSG00000042980	ADAM28
ENSG00000165023	DIRAS2
ENSG00000186297	GABRA5
ENSG00000129596	CDO1
ENSG00000162877	PM20D1
ENSG00000182240	BACE2
ENSG00000169169	CPT1C
ENSG00000234380	AP000330.8
ENSG00000139910	NOVA1
ENSG00000197632	SERPINB2
ENSG00000178919	FOXE1
ENSG00000255020	AF131216.5
ENSG00000251003	RP11-152P17.2
ENSG00000226278	PSPHP1
ENSG00000015475	BID
ENSG00000197134	ZNF257
ENSG00000164105	SAP30
ENSG00000174669	SLC29A2

ENSG00000232040	SCAND3
ENSG00000131941	RHPN2
ENSG00000233328	PFN1P1
ENSG00000230795	HLA-K
ENSG00000102755	FLT1
ENSG00000197046	SIGLEC15
ENSG00000006757	PNPLA4
ENSG00000189060	H1F0
ENSG00000237161	RP11-32B5.1
ENSG00000176659	C20orf197
ENSG00000154102	C16orf74
ENSG00000144488	ESPNL
ENSG00000048540	LMO3
ENSG00000075651	PLD1
ENSG00000260409	RP11-403B2.7
ENSG00000126562	WNK4
ENSG00000152527	PLEKHH2
ENSG00000272870	RP11-798M19.6
ENSG00000205045	SLFN12L
ENSG00000262576	PCDHGA4
ENSG00000114541	FRMD4B
ENSG00000248874	C5orf17
ENSG00000247095	MIR210HG
ENSG00000155511	GRIA1
ENSG00000143847	PPFIA4
ENSG00000099984	GSTT2
ENSG00000099974	DDTL
ENSG00000231007	CDC20P1
ENSG00000147862	NFIB
ENSG00000272787	KB-226F1.2
ENSG00000261613	RP11-20I23.13
ENSG00000234449	RP11-706O15.3
ENSG00000254913	RP4-791M13.5
ENSG00000158825	CDA

Supplementary Table 3.5- *C. elegans* strains used in the study

Strain	source	genotype
N2	CGC	
RB2260	CGC	*alfa-1(ok3062)*
EG1285	CGC	*unc-47p::GFP + lin-15(+)*
NL3321	CGC	*sid-1(pk3321)*
WM27	CGC	*rde-1(ne219)*
VP303	CGC	*rde-1(ne219) V; kbIs7. kbIs7 [nhx-2p::rde-1 + rol-6(su1006)*
TU3401	CGC	*(myo-2p::mCherry) + unc-119p::sid-1*
NR350	CGC	*rde-1(ne219) V; kzIs20 kzIs20[pDM#715(hlh-1p::rde-1) + pTG95(sur-5p::nls::GFP)].*
XE1375	CGC	*[unc-47p::mCherry] I. wpSi1 [unc-47p::rde-1::SL2::sid-1 + Cbr-unc-119(+)]*

DISCUSSION

INTRODUCTION

SUMMARY OF WORK PRESENTED

The work presented here has used *C. elegans* models to learn more about the toxicity of *C9orf72*. The studies that were conducted were designed to address the different modes of toxicity hypothesized. Hence, both loss of function and gain of RNA function models were examined and the two mechanisms are toxic to the *C. elegans* nervous system.

Using these models, we conducted a drug screen and identified the first molecules that could alleviate the toxicity of both *C9orf72* loss of function and gain of RNA function. The drugs identified were specific to *C9orf72* toxicity in worms, having little overlap with molecules that alleviate toxicity in other neuronal dysfunction and/or neurodegeneration models. Finally, a number of neuroprotective molecules were tested in *C9orf72* patient fibroblasts and some were shown to correct the dysregulated RNA expression observed in those cells.

It has been speculated that links between *C9orf72* and other ALS genes might help to identify toxic mechanisms that could lead to motor neuron death. To further understand the role of *C9orf72* across common biological pathways shared between different ALS genes, our laboratory is working toward the development of a large genetic

interaction map [2]. If genes interact genetically, it means that they function in the same or compensating cellular pathways [240]. Using different loss of function and transgenic *C. elegans* models, we have identified genes that can modify each other phenotypes. Using the loss of function model of *alfa-1/ C9orf72*, *C9orf72* was shown to interact differently with TDP-43 and FUS proteotoxicity where a decreased expression of *alfa-1/ C9orf72* exacerbated only TDP-43 toxicity.

In summary, the work presented here has led to the development of new animal models of *C9orf72*, increased our understanding regarding *C9orf72* toxicity and developed a new strategy to identify molecules that could alleviate *C9orf72* toxicity. Therefore, these three topics will be discussed in more details in this chapter.

ANIMAL MODELS OF *C9ORF72*

Since the beginning of our work, many new animal models were developed to better understand *C9orf72*. Interestingly, even though none of these models had been published at the time we undertook the development of our *C. elegans* models, many characteristics are shared between our models and these. Here is a brief summary of these models followed by a discussion on how they relate to our *C. elegans* models.

SUMMARY OF NEW C9ORF72 MODELS

The first published *C9orf72* model organism used *Drosophila*. *C9orf72* is not conserved in *Drosophila* but by expressing different GGGGCC repeat lengths under various promoters, many groups have worked to gain insight into *C9orf72* toxicity [122,124,241-244]. Interestingly in many of these models only 30 GGGGCC repeats were required to induce toxicity in the fly eyes and nervous system, leading to age-dependent neuromuscular problems, motility impairment, eye structure disorganization and lethality [124,241,242]. It is also noteworthy that even if the production of *C9orf72*-derived dipeptide aggregates and RNA foci have been observed in several models, a direct causative link between these pathological features and toxicity could not always be observed.

In a first attempt to develop a mouse model to understand *C9orf72* toxicity, Chew *et al*. used an adenoviral expression system to express a pathogenic GGGGCC repeat in the central nervous system. Six months after a single injection of the virus, RNA foci, dipeptides and aggregates of phosphorylated TDP-43 were observed in post-mortem tissues. Neither neuronal loss nor motility problem was observed, but the mice had slight balance problem and an anxiety behaviour [245]. Interestingly, these observations are

supported by the recent reports of two independent groups that have also developed *C9orf72* mouse models but failed to observe neurodegeneration [246,247]. Both groups used a bacterial artificial chromosome (BAC) system to express a partial (exons 1-6) or full length human *C9orf72* that comprised a pathogenic GGGGCC repeat (ranging from 100-1,000 units long). Overall, mouse model data have revealed that even if RNA foci and different dipeptides were produced in glial cells and neurons, neurodegeneration and motility impairments were not observed, even at advanced age, in the animals [246,247]. Also, it appears that animals did not openly exhibit cognitive features similar to other FTD mouse models [246,247]. Taken together, these results suggest that neither cognitive deficit nor motor deficit could be caused only by the expression of human *C9orf72* harbouring a pathogenic repeat in mice.

CONCLUSION

Each model organism comes with its own inherent advantages and disadvantages. For the models developed in this work, the use of the RNAi system to expose neurons to a RNA containing a pathogenic GGGGCC repeat expansion proved to be useful; albeit it could not insure an equal expression level in all individual animals. While it is known that loss of *alfa-1* expression causes a significant motor phenotype in worm, we cannot exclude that potential mutations in other genes participate to this phenotype. Moreover, most *C9orf72* positive patients are heterozygous and as such only have a partial loss of function, a condition that could not be evaluated in the *alfa-1* loss of function model. However, the models developed in this work share common characteristics with other *C. elegans* models of ALS and with other repeat disorder models.

Published work shows that *C. elegans* transgenic models of TDP-43 exhibit specific neuronal vulnerability of the GABAergic neurons[198], similar to what is observed in the *alfa-1(ok3062)* models. This interesting aspect shows that in worm some neurons are more vulnerable to ALS proteins, recapitulating an important feature of ALS pathogenesis in humans. Also, our data have demonstrated that *alfa-1/ C9orf72* was involved in osmotic stress response which was also the case for *fust-1/FUS* and *tdp-1/TDP-43* [205](Therrien et al. annexe I.) In different polyglutamine models, impaired osmotic stress response was shown to be involved in protein aggregation [248]. Therefore, knowing the importance of aggregation in ALS pathogenesis (see section *Pathological characteristics of ALS*), the involvement of three major ALS proteins linked to the osmotic stress response could highlight a specific major pathogenic pathway. Osmolarity in the human nervous system is one of the most regulated processes and do not change much in normal state[249], therefore it is difficult to evaluate the importance of those findings in regards to ALS pathogenesis. However, these data suggest that cells exhibiting a loss of expression of *alfa-1/ C9orf72*, *tdp-1/TDP-43* or *fust-1/FUS* might be more prone to protein aggregation leading to increased neuronal vulnerability.

The work presented here has also shown that expression of a pathogenic GGGGCC repeat results in neuronal phenotypes in *C. elegans*. The toxicity we observed is similar to what other groups have observed in *Drosophila*, but different from the mammalian models published so far. It would not be the first time that simple model organisms recapitulate ALS better than mouse models from which many failed to fully recapitulate ALS pathogenesis (see section *Current models to understand ALS*). It is important to note that many repeat containing genes forming aggregates are more toxic

in their cleaved form than in the full-length version (among others [250,251]). Also, Huntington's disease mouse models have shown that expression of a truncated version of the *HTT* gene is more toxic than the expression of a full-length gene where the truncated version caused earlier and more severe phenotypes in the animals (reviewed by Pouladi *et al.* [252]). Thus, it is not surprising that overexpression of the GGGGCC repeat by itself is toxic in many model organisms, and that models based on BAC expression might exhibit reduced toxicity. Therefore our work and the work of others suggest that the surrounding region of the GGGGCC repeat of *C9orf72* is highly important and could influence its toxicity.

Simple model organisms can be useful but cannot always model complex compensatory mechanisms that are observed in humans. For example, *FUS*, *EWS* and *TAF15* genes are highly similar RNA binding proteins that are genetically linked to ALS [20,21,253,254]. In *C. elegans*, the deletion of *fust-1*, the orthologue of all three genes, causes age-dependent neurodegeneration (Therrien *et al.* annexe I). However, in mice, decreased expression of *FUS* causes an increased expression of *TAF15* and *EWS* [255], therefore suggesting a complex relationship between these proteins. These data suggest that this family of proteins play an essential and complex role in maintaining neuronal integrity. In light of data from us and other groups that have tried to develop *C9orf72* models, we speculate that *C9orf72* and its GGGGCC repeat play a role in neuronal integrity but compensatory mechanisms may alleviate their effect in mammalian systems.

With the rapid advances of genetic technologies, new models can be developed to better recapitulate the conditions that are observed in patients. Considering that no sole model organism can fully recapitulate the cellular physiology and structure of humans, it is important to take into account the different models where a relevant phenotype provides useful insight to advance ALS research.

Small pathogenic repeat lengths of GGGGCC seem to be highly toxic in model organisms (*i.e.* only 30 repeats are required in *Drosophila* to be toxic and 50 units in *C. elegans*) but in patients the repeat length seems to be much longer. The blurry characterization of the length of the GGGGCC repeat of *C9orf72* in ALS and/or FTD patients hampers efforts to develop relevant models. The most toxic length of the GGGGCC repeat needs to be identified to avoid the development of non-toxic GGGGCC models.

The next generation of *C9orf72 C. elegans* models should take into account both the decreased expression of *alfa-1/ C9orf72* and the toxicity of the pathogenic GGGGCC repeat. New technologies such as CRISPR/Cas9 would allow the insertion of a pathogenic GGGGCC repeat within the first intron of *alfa-1* gene, and is currently under development in the laboratory. Also, using a wide variety of repeat lengths within the same cells, along with similar expression level, could help elucidate the question regarding the toxicity of the repeat length.

C. elegans might not fully recapitulates ALS pathogenesis observed in patients but, its ease of cultivation and its speed make it a great model to study specific aspects of *C9orf72* pathogenesis. The current models, as well as newer ones could be used to

identify modifiers of *alfa-1/ C9orf72* expression, or modifiers of translation mechanisms that induce the production of toxic dipeptide species.

MODES OF TOXICITY OF *C9ORF72*

Since the beginning of this project, many research articles have been published regarding the potential toxicity caused by the expanded GGGGCC repeat of *C9orf72*. Here is a brief summary including the most important findings regarding this ongoing debate and how the work presented here played an important role in this debate.

SUMMARY OF ADVANCEMENTS OF RESEARCH

LOSS OF *C9ORF72* EXPRESSION

Loss of expression of *C9orf72* has been observed at the RNA and protein levels in different regions of the brain and spinal cord [47,117]. Similar to what was shown in this work using *C. elegans*, in zebrafish, complete loss of expression of the orthologue of *C9orf72*, *zC9orf72*, was shown to induce motility problems that could be rescued by expression of the human *C9orf72* [256], suggesting that both genes play similar function in neurons. However, in mice, transient decreased expression of *C9orf72* in the central nervous system and a conditional knock-out model, did not induce pathological (TDP-43 and p62 aggregates) or behavioural changes in the animals [238] [112]. Also, some groups have suggested that hypermethylation might be neuroprotective[257,258], however, these data still remain to be confirmed in neuronal cells.

RAN TRANSLATION LEADING TO PRODUCTION OF DIPEPTIDES

As observed in other repeat disorders, even though the GGGGCC repeat of *C9orf72* is found in the first intron of the gene, it was speculated that it could cause the production of dipeptides repeat proteins through RAN translation (for more details see *Chapter 3-Introduction*). The presence of dipeptides was shown in numerous patients neuronal tissues including the cerebellum, spinal cord and frontal cortex [236,259,260]. The glycine-alanine and glycine-proline seem to be the most abundant dipeptides produced [259,260] and glycine-proline was even found in cerebrospinal fluid of *C9orf72* positive ALS patients [261].

Using different model organisms, the toxicity of the different dipeptides was assessed. Results of these studies are, however, controversial and expression level and tissue specificity could explain some variations between the different models. The dipeptides that seem to be the most toxic are the glycine-arginine and proline-arginine dipeptides [122,169,242,262]. However, these are among the least abundant dipeptides found in patient post-mortem tissues [259,260,263]. Also, toxicity of the dipeptides produced by other repeat disorder transcripts (such as *Atx-8* and *Htt* transcripts) is also unclear. Therefore, additional studies are required to evaluate the importance of dipeptide formation in ALS pathogenesis.

RNA FOCI

RNA foci were observed in *C9orf72* positive patients neuronal and non-neuronal cells [259]. The RNA aggregates were shown to bind to numerous RNA binding protein *in vitro* [264] and to bind and sequester some of these RNA binding proteins in the nucleus [226,265]. PUR-

ALPHA was identified by many groups as one of the main binding protein of the GGGGCC expanded RNA, and was found in the RNA foci of *in vitro* and *in vivo* models of *C9orf72* as well as in patients derived cells [115,127,241]. The overexpression of PUR-ALPHA rescued phenotypes caused by the expression of GGGGCC repeat in *Drosophila* and in transfected cells [241]. Interestingly, PUR-ALPHA was also found to bind the CGC repeat of *FMR1* causing Fragile-X ataxia/tremor syndrome. PUR-ALPHA is an RNA binding protein that binds purine rich sequence and a knock-out mouse model has shown that it is involved in neuronal proliferation and postnatal brain development[266]. Therefore these data suggest a general implication for this protein within C-G rich repeat disorders and in the nervous

system. However, so far the data have not shown a clear correlation between the presence of RNA foci and neuronal toxicity.

Conclusion

Loss of function of the protein affected by the repeat, toxicity of the RNA foci, and the production of toxic dipeptides have all been shown to occur in many repeat disorders (for more details see *Repeat disorders* section). Generally, it is thought that expanded repeats employ several mechanisms to ultimately trigger neurodegeneration. The data presented here suggest that both a loss of function, and RNA toxicity, that could lead to

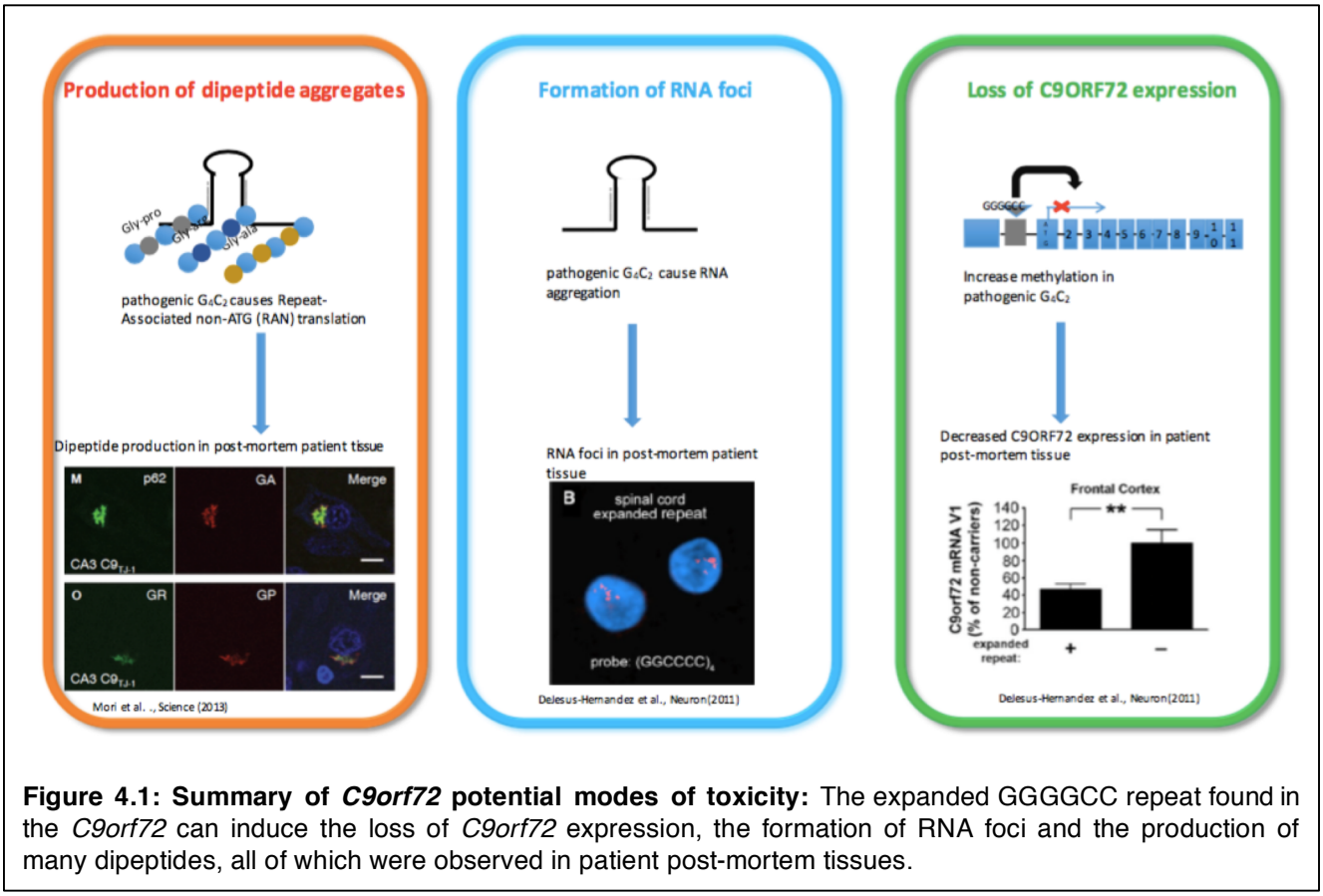


Figure 4.1: Summary of *C9orf72* potential modes of toxicity: The expanded GGGGCC repeat found in the *C9orf72* can induce the loss of *C9orf72* expression, the formation of RNA foci and the production of many dipeptides, all of which were observed in patient post-mortem tissues.

the formation of RNA foci or to the production of dipeptides, are toxic to neurons.

We have shown that a constitutive loss of *alfa-1/ C9orf72* function is detrimental to neurons. However, mouse models using a transient or conditional loss of expression did not cause any neuronal abnormalities in the animals. Expression level of *C9orf72* throughout the progression of the disease is not well characterized, but the absence of *C9orf72* during development could lead to a loss of neuronal integrity later in life. Furthermore, in most neurodegenerative disorders caused by a genetic variant, the mutant protein is present and expressed from birth but only become toxic during aging.

The work presented here has shown that the expression of RNA containing an expanded GGGGCC repeat is toxic to neurons and could lead to abnormal RNA expression of many other transcripts. Also, other groups have shown that in patient cells, where a decreased expression of *C9orf72* is not observed, abnormal RNA expression is seen[226]. Several research groups have tried to differentiate between the toxicity of the RNA and the dipeptides and all have concluded that the GGGGCC repeat was more toxic at the protein level than at the RNA level [242]. However, no clear validation has been done in patient tissue where production of dipeptides does not correlate with signs of neuronal degeneration [263,267]. It is also noteworthy to remember that the dipeptides form aggregates, and for most neurodegenerative disorders, the toxicity of these protein aggregates is still unclear. Our data do not allow us to suggest if the neurotoxicity is caused by the presence of dipeptides and/or the RNA foci, but the sole presence of the RNA containing a pathogenic GGGGCC repeat affects neurons. However, in patient-derived cells, where many abnormal RNA expression events were observed and could impair cellular function, only the dipeptides were observed.

The work presented here also suggests that tissue specificity is important for *C9orf72* toxicity, where a loss of expression is detrimental to many cell types in *C. elegans* while only neurons are affected by the GGGGCC repeat. These results are novel and bring a new perspective about *C9orf72* toxicity because most expression models published so far only characterize the impact of the expanded GGGGCC in a single type of cells or neurons. Tissue specificity is an important aspect of ALS pathogenesis and it is not surprising that change cell vulnerability is observed. The best example of tissue specificity regarding *C9orf72* toxicity is the decreased expression observed in many neuronal and blood cells but not in fibroblasts. Different groups, including ourselves, have used fibroblast cell lines derived from patients and could not observe a decrease in *C9orf72* expression [114,226]. Therefore, it shows that fibroblasts may be useful to understand RAN translation leading to the production of dipeptides or RNA foci toxicity, but we cannot assume that they will give full insights into neuronal characteristics of *C9orf72*. Most ALS proteins are ubiquitously expressed but neuronal loss is focused, therefore neuronal specificity is an important, unresolved topic relevant to ALS.

For many repeat disorders, where it was first speculated to be a toxic gain of function mechanism, studies have shown that a decreased expression of the endogenous gene could contribute to neuronal toxicity [268-270]. Thus, there may be a cellular environment where both a gain, and a loss of function participate to the pathogenesis. Also, the CGC repeat found in *FMR1* provides a good example of a repeat causing a loss of function and a gain of function toxicity [151]. When the repeat is between 70-200 units, it causes an increase in mRNA expression which leads to a neurodegenerative disorder called Fragile-X tremor/ataxia syndrome and the toxicity is speculated to be a gain of function

of the protein. When the repeat is longer than 200 units, it causes a decreased in *FMR1* gene expression leading to a neurodevelopmental disorder called Fragile-X syndrome, likely causing toxicity through the loss of *FMR1* function (reviewed in *Fragile X syndrome and toxicity of a decreased gene expression* section). Therefore, a similar mechanism could participate to explain the involvement of *C9orf72* in ALS and FTD.

It is important to note that even though mutations in *TARDBP* and *FUS* are speculated to cause a gain of function, loss of the normal functions of the proteins they encode have been shown to be detrimental to cells [128]. Similar mechanisms could easily be envisaged for *C9orf72* toxicity where loss of expression, and gain of toxic properties from the RNA leave the neurons vulnerable to cell death. For example, if *C9orf72* is confirmed to act with RAB proteins and affect protein degradation, a decreased expression of *C9orf72* could exacerbate the toxicity of the dipeptides by impairing their proper degradation.

PERSPECTIVES

Some characteristics of *C9orf72* toxicity could help us respond to the questions related to *C9orf72* toxicity, however. What is the timeline of these phenomena? What happens first to initiate the toxicity cascade that ultimately kills neurons? Patient tissues are just a snapshot of the histopathological view when the patients die. The RNA foci and dipeptide aggregates could be involved in protecting neurons, since the neurons observed are the ones that survived. This question is highly relevant since mouse models exhibiting RNA foci and dipeptide aggregates caused by the expression of a pathogenic GGGGCC repeat do not show any neuronal toxicity [245-247].

Are all types of ALS similar? Could a better characterization of the symptoms observed in patients lead to a categorization of different ALS cases? If we speculate that *C9orf72* could cause a dose-dependent effect similar to what is observed in Fragile-X syndrome/tremor ataxia, a comprehensive characterization of patient phenotypes will be necessary. Finally, a clear connection between *C9orf72* GGGGCC repeat length, *C9orf72* expression level and motor and cognitive dysfunctions will be necessary. Interestingly, here, model organisms could prove to be highly useful to evaluate the variation of *C9orf72* expression, the formation of RNA foci and production of dipeptides throughout development of the disease.

ALS DRUG DEVELOPMENT

Using different C. *elegans* models, we employed a multistep strategy to identify drugs that could alleviate *C9orf72* toxicity and could highlight potential therapeutic pathways for *C9orf72* positive patients. We identified small molecules that alleviated the toxicity from the loss of *alfa-1/ C9orf72* expression, the GGGGCC RNA toxicity in *C. elegans,* and attenuated the abnormal RNA expression observed in patient-derived cells. Here is a summary of the standard drug development process and the challenges posed to the identification of effective ALS therapies.

USUAL DRUG DEVELOPMENT PROCESS[2]

Drug development is a lengthy process that can often take two decades before a drug makes it to the market. According to the US Food and Drug Administration (FDA), molecules are first identified in the research laboratory and are then tested in different *in vitro* and *in vivo* models (called preclinical trials). These models are often mouse models recapitulating key features of the disease. If the molecule is efficient, and

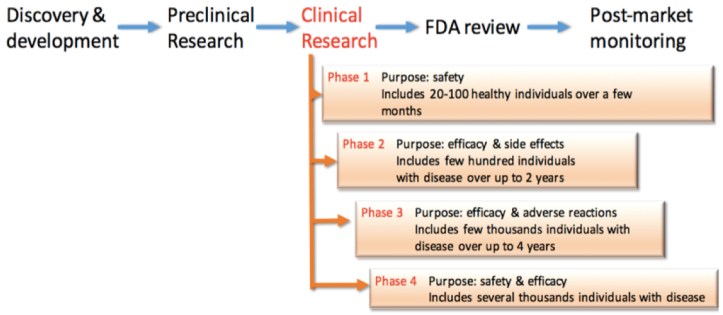


Figure 4.2: Outline of drug development

[2] FDA website http://www.fda.gov/ForPatients/Approvals/default.htm page visited December 13, 2015

if they are available, the molecules are tested in larger mammalian models. From there, clinical trials are started that include four phases, testing an increasing number of individuals with or without a disease to assay for drug safety, efficacy and potential side effects (Figure 4.2). Finally, drugs will be evaluated by the health agency of the country, and if the agency finds it to be safe, the drug will be released for treatment of ALS

Needless to say, drug development is a lengthy and expensive process. Using *C. elegans* and FDA approved molecules, we speculated that we could develop a high-throughput drug screening of molecules that would specifically target some specific phenotypes in the worm. Development of new *C. elegans* models is relatively fast and repurposing molecules that are already approved by the FDA could accelerate their testing in clinical settings.

CHALLENGES FOR THE DEVELOPMENT OF ALS DRUGS

EFFICACY OF RILUZOLE

Riluzole was first approved in 1995 and still remains the only drug approved to treat ALS patients. It has a modest effect, increasing survival from 3-6 months[6]. The exact function of riluzole is still unclear but it is speculated to act on excitotoxicity. The first evaluation of riluzole in ALS patients was published in 1994 and the rationale for it was solely based on patient characteristics which include an observed decreased in glutamate uptake and abnormal glutamate metabolism in patients[271]. Riluzole has also been shown to improve phenotypes in Spinal muscular atrophy mouse [272], another motor neuron disorder, and

clinical trials are now being done to test its efficacy in patients[3]. Riluzole is also being tested for many neurodegenerative disorders including Parkinson's disease, Alzheimer's disease, Huntington's disease and cerebellar ataxia, and for treatment in many psychiatric disorders such as bipolar disorder, obsessive-compulsive disorder, post-traumatic stress injury and depression[2]. Thus riluzole has general activity in restoring neuronal function.

THE DIFFICULTIES IN ALS CLINICAL TRIALS

For the past 50 years, over 50 clinical trials were undertaken to test molecules that could treat or delay symptoms in ALS patients[166]. Even now, more than 40 clinical trials are being conducted worldwide to test the efficacy of molecules or care management that could alleviate disease progression[4] . In just the past 10 years, more than 15 molecules have failed during phases 2 or 3 of clinical trial [166]. Many of these molecules were believed to target mechanisms with central roles in ALS pathogenesis including; excitotoxicity, neuroinflammation, oxidative stress, apoptosis, mitochondrial dysfunction, autophagy and astrocyte functions.

Many researchers are questioning the failures of these clinical trials suggesting that the simple measurement of lifespan extension in patients is inappropriate for a

[3] U.S. National Institutes of Health-Clinical trials website clinicaltrials.gov page visited Dec 26, 2015

[4]U.S. National Institutes of Health-Clinical trials website https://www.clinicaltrials.gov/ct2/results?cond=ALS&pg=2 page visited December 26, 2015

disease in which the disease cause, duration and development is so variable [273]. However, the lack of efficient biomarkers to predict disease progression had made the use of other endpoints very difficult [274] .

The use of mouse models in pre-clinical studies has also been questioned. If molecules have made it to phase 2 and 3, it is because they were successful in pre-clinical phases using mouse models. To this day, the transition of observation that was made with mouse models to humans has been disappointing, thus it is difficult to predict the effectiveness of molecules that are promising in mice to ALS patients.

CONCLUSION

An important question is how to identify drugs that could delay the progression of ALS? Although bolstered by promising results in model systems, drugs that have targeted known pathogenic mechanisms have failed so far in ALS clinical trials. Additionally, even with the knowledge that riluzole has a beneficial effect, the ALS community has been unable to make this drug more effective more than 20 years after its approval. With the fast development of genetic techniques, many genes have been linked to ALS, and while some share common functions, many questions remain regarding how they cause motor neuron death. Therefore, finding a single molecule that can target one specific toxic mechanism and improve survival in most ALS patients is becoming seemingly impossible. Also, ALS has become a spectrum of motor neuron disorders that can or cannot include cognitive deficits, increasing the complexity in finding a single molecule

to address all causes. We speculated that identifying specific phenotypes in a model organism caused by a specific genetic variant would be the most promising approach.

Molecules have typically been developed based on their known target and the target pathway that is believed to be involved in the cellular toxicity. However, some groups have speculated that phenotypic screens could be more efficient than target-based screen to identify molecules that could be relevant in human[275,276]. We speculated in this project that *C. elegans* is a valid model to carry large molecule screens based on a specific phenotype, independently of the known function of the molecules. Furthermore, testing the chosen molecules in patient derived-cells has allowed us to confirm that some of these molecules could also act on pathological features observed in *C9orf72* positive patients. Using FDA approved molecules has helped identify bioactive molecules, many of which are known to cross the blood-brain barrier and increasing their potential as therapeutics for ALS patients. Lastly, similarly to riluzole, many of the drugs identified here are known to be beneficial in psychiatric diseases, demonstrating their ability to influence neuronal phenotypes.

The work presented here is focused on *C. elegans* models of *C9orf72* and we present molecules that specifically rescue the neuronal loss in an *in vivo* context Furthermore, we have shown that the drugs identified in worms are also active in mammalian systems, here by reducing the abnormal RNA expression observed in patient-derived cells. Of course, optimisation of these drugs might be necessary before they are tested in humans but they are prime candidates for testing in other *C9orf72* models. Some of these molecules have many cellular targets but finding which of these

targets are involved in the neuroprotection against *C9orf72* toxicity could highlight specific pathways that are protective against *C9orf72* toxicity, therefore opening new research avenues regarding *C9orf72* research.

PERSPECTIVES

Gene therapy and stem cell transplants represent alternatives to the use of small-molecule for the treatment of ALS. The identification of CRISPR/Cas9 method to modify genetic variant is promising and much effort is being done to investigate this technique in human[277-279]. The reversal of the toxic genetic variant would be the most upstream treatment for patients in which ALS as a genetic composition. However, ethical and safety considerations will delay its application in humans.

Personalized medicine is becoming useful and popular, so a similar mechanism could be proposed for the future of ALS drug discovery. Therefore, finding molecules that would specifically act on *C9orf72* would be an important first step toward the treatment of ALS. In the best-case scenario, finding drugs that could completely halt neuronal death would be ideal. However, the identification of molecules that could slow disease progression in one type of ALS would be a major step for patients. Given that aging is a major contributing factor for diseases like ALS, delaying disease onset may effectively cure the disease in some cases.

The molecules identified during this project should be tested in other models of *C9orf72* exhibiting neuronal loss, either mice or iPSC-neurons derived from patients.

Also, finding the targets of those drugs could confirm their potential actions and open new research avenues regarding *C9orf72* toxicity.

CONCLUSION

At this point on, 10-20% of the genetic mutations underlying of ALS cases are known and many more genetic variations and risk factors remain to be identified. Among the genetic risk factors, the *C9orf72* GGGGCC repeat expansion is currently the most prevalent known cause of fALS and sALS. Even though much effort is still needed to understand the cases that cannot be explained by genetics, identification of molecules that could alleviate the toxicity of, at least one toxic protein would have a tremendous impact. The work presented here aimed to develop new models to understand *C9orf72* toxicity and to identify molecules that could alleviate neurodegeneration.

Much remains to be understood about the toxicity of *C9orf72*. The repeat is subject to methylation causing a decreased expression of *C9orf72*. Additionally, RNA containing the expanded repeat forms nuclear RNA foci, and has been shown to induce the production of several dipeptides. To assess all of these modes of toxicity we have developed and characterized two new *C. elegans* models. Our worm models helped us better understand about the loss of function of *C9orf72,* and the toxicity of the RNA containing the expanded repeat. Both modes of toxicity were detrimental to neurons although they seem to have distinct cellular toxicity mechanisms. Using these models, we identified molecules that suppressed the toxicity in both models. Finally, we have tested the effectiveness of these molecules in ALS using *C9orf72* positive patient cells.

The work presented here includes some of the first models to study *C9orf72* in which neuronal integrity was evaluated. Also, this work represents the first *C9orf72* specific drug screening.

The field of *C9orf72* research is still young but is beginning to uncover common themes between all neurodegenerative disorders. Study of *C9orf72* shows the similarities between ALS and many other diseases including age-related diseases, motor diseases, repeat disorders, and any disorder caused by abnormal RNA and protein homeostasis. Therefore, the work and models presented here will hopefully be proven useful not only for ALS patients, but also for other age-related disorders, motor neuron disease, repeat disorders and any diseases where RNA and/or protein homeostasis are involved.

References

1. Therrien, M. & Parker, J. A. Worming forward: amyotrophic lateral sclerosis toxicity mechanisms and genetic interactions in Caenorhabditis elegans. *Front Genet* **5,** 1–13 (2014).
2. Therrien, M. & Parker, J. A. Deciphering genetic interactions between ALS genes using C. elegans. *worm* **3,** e29047 (2014).
3. Boylan, K. Familial Amyotrophic Lateral Sclerosis. *Neurol Clin* **33,** 807–830 (2015).
4. Andersen, P. M. & Al-Chalabi, A. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nat Rev Neurol* **7,** 603–615 (2011).
5. Byrne, S. *et al.* Proposed criteria for familial amyotrophic lateral sclerosis. *Amyotroph Lateral Scler* **12,** 157–159 (2011).
6. Kiernan, M. C. *et al.* Amyotrophic lateral sclerosis. *The Lancet* **377,** 942–955 (2011).
7. Turner, M. R. *et al.* Controversies and priorities in amyotrophic lateral sclerosis. *The Lancet Neurology* **12,** 310–322 (2013).
8. Traynor, B. J. *et al.* Clinical features of amyotrophic lateral sclerosis according to the El Escorial and Airlie House diagnostic criteria: A population-based study. *Arch Neurol* **57,** 1171–1176 (2000).
9. Al-Chalabi, A. & Visscher, P. M. Motor neuron disease: Common genetic variants and the heritability of ALS. *Nat Rev Neurol* **10,** 549–550 (2014).
10. Guerreiro, R., Brás, J. & Hardy, J. SnapShot: Genetics of ALS and FTD. *Cell* **160,** 798–798.e1 (2015).
11. Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362,** 59–62 (1993).
12. Bowling, A. C., Schulz, J. B., Brown, R. H. & Beal, M. F. Superoxide dismutase activity, oxidative damage, and mitochondrial energy metabolism in familial and sporadic amyotrophic lateral sclerosis. *Journal of Neurochemistry* **61,** 2322–2325 (1993).
13. Al-Chalabi, A. *et al.* The genetics and neuropathology of amyotrophic lateral sclerosis. *Acta Neuropathol* **124,** 339–352 (2012).
14. Pickles, S. & Vande Velde, C. Misfolded SOD1 and ALS: zeroing in on mitochondria. *Amyotroph Lateral Scler* **13,** 333–340 (2012).
15. Tafuri, F., Ronchi, D., Magri, F., Comi, G. P. & Corti, S. SOD1 misplacing and mitochondrial dysfunction in amyotrophic lateral sclerosis pathogenesis. *Front Cell Neurosci* **9,** 336 (2015).
16. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314,** 130–133 (2006).

17. Kabashi, E. *et al.* TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nat. Genet.* **40,** 572–574 (2008).
18. Sreedharan, J. *et al.* TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* **319,** 1668–1672 (2008).
19. Lagier-Tourenne, C., Polymenidou, M. & Cleveland, D. W. TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum. Mol. Genet.* **19,** R46–R64 (2010).
20. Kwiatkowski, T. J. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323,** 1205–1208 (2009).
21. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323,** 1208–1211 (2009).
22. Fontana, F., Siva, K. & Denti, M. A. A network of RNA and protein interactions in Fronto Temporal Dementia. *Front Mol Neurosci* **8,** 9 (2015).
23. Leblond, C. S., Kaneb, H. M., Dion, P. A. & Rouleau, G. A. Dissection of genetic factors associated with amyotrophic lateral sclerosis. *Experimental Neurology* **262 Pt B,** 91–101 (2014).
24. Renton, A. E., Chiò, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nat Neurosci* **17,** 17–23 (2014).
25. Lattante, S. *et al.* Contribution of major amyotrophic lateral sclerosis genes to the etiology of sporadic disease. *Neurology* **79,** 66–72 (2012).
26. Al-Chalabi, A. & Hardiman, O. The epidemiology of ALS: a conspiracy of genes, environment and time. *Nat Rev Neurol* **9,** 617–628 (2013).
27. Veltman, J. A. & Brunner, H. G. De novo mutations in human genetic disease. *Nat. Rev. Genet.* **13,** 565–575 (2012).
28. Girard, S. L. *et al.* Increased exonic de novo mutation rate in individuals with schizophrenia. *Nat. Genet.* **43,** 860–863 (2011).
29. Alexander, M. D. *et al.* 'True' sporadic ALS associated with a novel SOD-1 mutation. *Ann Neurol.* **52,** 680–683 (2002).
30. Laffita-Mesa, J. M. *et al.* De novo mutations in ataxin-2 gene and ALS risk. *PLoS ONE* **8,** e70560 (2013).
31. Calvo, A. *et al.* A de novo nonsense mutation of the FUS gene in an apparently familial amyotrophic lateral sclerosis case. *Neurobiology of Aging* **35,** 1513.e7–11 (2014).
32. Chiò, A. *et al.* A de novo missense mutation of the FUS gene in a 'true' sporadic ALS case. *Neurobiology of Aging* **32,** 553.e23–6 (2011).
33. DeJesus-Hernandez, M. *et al.* De novo truncating FUS gene mutation as a cause of sporadic amyotrophic lateral sclerosis. *Hum. Mutat.* **31,** E1377–89 (2010).
34. Hübers, A. *et al.* De novo FUS mutations are the most frequent genetic cause in early-onset German ALS patients. *Neurobiology of Aging* **36,** 3117.e1–3117.e6 (2015).
35. Mackenzie, I. R., Rademakers, R. & Neumann, M. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *Lancet Neurol* **9,** 995–1007 (2010).

36. Clement, A. M. *et al.* Wild-type nonneuronal cells extend survival of SOD1 mutant motor neurons in ALS mice. *Science* **302,** 113–117 (2003).
37. Boillée, S. *et al.* Onset and progression in inherited ALS determined by motor neurons and microglia. *Science* **312,** 1389–1392 (2006).
38. Haidet-Phillips, A. M. *et al.* Altered astrocytic expression of TDP-43 does not influence motor neuron survival. *Experimental Neurology* **250,** 250–259 (2013).
39. Serio, A. *et al.* Astrocyte pathology and the absence of non-cell autonomy in an induced pluripotent stem cell model of TDP-43 proteinopathy. *Proceedings of the National Academy of Sciences* **110,** 4697–4702 (2013).
40. Schipper, L. J. *et al.* Prevalence of brain and spinal cord inclusions, including dipeptide repeat proteins, in patients with the C9ORF72 hexanucleotide repeat expansion: A systematic neuropathological review. *Neuropathology and Applied Neurobiology* n/a–n/a (2015). doi:10.1111/nan.12284
41. Ling, S.-C. *et al.* ALS-associated mutations in TDP-43 increase its stability and promote TDP-43 complexes with FUS/TLS. *Proceedings of the National Academy of Sciences* **107,** 13318–13323 (2010).
42. Majcher, V., Goode, A., James, V. & Layfield, R. Autophagy receptor defects and ALS-FTLD. *Mol. Cell. Neurosci.* **66,** 43–52 (2015).
43. Warren, J. D., Rohrer, J. D. & Rossor, M. N. Clinical review. Frontotemporal dementia. *BMJ* **347,** f4827 (2013).
44. Ringholz, G. M. *et al.* Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* **65,** 586–590 (2005).
45. Woolley, S. C. & Strong, M. J. Frontotemporal Dysfunction and Dementia in Amyotrophic Lateral Sclerosis. *Neurol Clin* **33,** 787–805 (2015).
46. Burrell, J. R., Kiernan, M. C., Vucic, S. & Hodges, J. R. Motor neuron dysfunction in frontotemporal dementia. *Brain* **134,** 2582–2594 (2011).
47. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* **72,** 245–256 (2011).
48. Renton, A. E. *et al.* A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* **72,** 257–268 (2011).
49. Deng, H.-X. *et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477,** 211–215 (2011).
50. Daoud, H. *et al.* C9orf72 hexanucleotide repeat expansions as the causative mutation for chromosome 9p21-linked amyotrophic lateral sclerosis and frontotemporal dementia. *Arch Neurol* **69,** 1159–1163 (2012).
51. Mori, F. *et al.* Valosin-containing protein immunoreactivity in tauopathies, synucleinopathies, polyglutamine diseases and intranuclear inclusion body disease. *Neuropathology* **33,** 637–644 (2013).
52. Woulfe, J., Gray, D. A. & Mackenzie, I. R. A. FUS-immunoreactive intranuclear inclusions in neurodegenerative disease. *Brain Pathol.* **20,** 589–597 (2010).
53. Kon, T. *et al.* ALS-associated protein FIG4 is localized in Pick and Lewy

bodies, and also neuronal nuclear inclusions, in polyglutamine and intranuclear inclusion body diseases. *Neuropathology* **34,** 19–26 (2014).
54. Doi, H. *et al.* RNA-binding protein TLS is a major nuclear aggregate-interacting protein in huntingtin exon 1 with expanded polyglutamine-expressing cells. *J. Biol. Chem.* **283,** 6489–6500 (2008).
55. Tan, C.-F. *et al.* Selective occurrence of TDP-43-immunoreactive inclusions in the lower motor neurons in Machado-Joseph disease. *Acta Neuropathol* **118,** 553–560 (2009).
56. Elden, A. C. *et al.* Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* **466,** 1069–1075 (2010).
57. Farg, M. A. *et al.* Ataxin-2 interacts with FUS and intermediate-length polyglutamine expansions enhance FUS-related pathology in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **22,** 717–728 (2013).
58. Pulst, S. M. *et al.* Moderate expansion of a normally biallelic trinucleotide repeat in spinocerebellar ataxia type 2. *Nat. Genet.* **14,** 269–276 (1996).
59. Aguzzi, A. & O'Connor, T. Protein aggregation diseases: pathogenicity and therapeutic perspectives. *Nat Rev Drug Discov* **9,** 237–248 (2010).
60. PhD, B. D. *et al.* Personal View Targeting α-synuclein for treatment of Parkinson's disease:mechanistic and therapeutic considerations. *The Lancet Neurology* **14,** 855–866 (2015).
61. Michalik, A. & Van Broeckhoven, C. Pathogenesis of polyglutamine disorders: aggregation revisited. *Hum. Mol. Genet.* **12 Spec No 2,** R173–86 (2003).
62. Matus, S., Valenzuela, V., Medinas, D. B. & Hetz, C. ER Dysfunction and Protein Folding Stress in ALS. *Int J Cell Biol* **2013,** 674751 (2013).
63. Vériène, J., Fossouo, L. & Parker, J. A. Neurodegeneration in C. elegans models of ALS requires TIR-1/Sarm1 immune pathway activation in neurons. *Nat Commun* **6,** 7319 (2015).
64. MD, P. M. T. H. *et al.* Neuroinflammation in Alzheimer's disease. *The Lancet Neurology* **14,** 388–405 (2015).
65. Hirsch, E. C. & Hunot, S. P. Neuroinflammation in Parkinson's disease: a target for neuroprotection? *The Lancet Neurology* **8,** 382–397 (2009).
66. Hoozemans, J. J. M. & Scheper, W. Endoplasmic reticulum: the unfolded protein response is tangled in neurodegeneration. *Int. J. Biochem. Cell Biol.* **44,** 1295–1298 (2012).
67. Mercado, G., Valdés, P. & Hetz, C. An ERcentric view of Parkinson's disease. *Trends Mol Med* **19,** 165–175 (2013).
68. Vance, C. *et al.* Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3. *Brain* **129,** 868–876 (2006).
69. Morita, M. *et al.* A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology* **66,** 839–844 (2006).
70. Valdmanis, P. N. *et al.* Three families with amyotrophic lateral sclerosis and frontotemporal dementia with evidence of linkage to chromosome 9p. *Arch Neurol* **64,** 240–245 (2007).

71. Gómez-Tortosa, E. *et al.* C9ORF72 hexanucleotide expansions of 20-22 repeats are associated with frontotemporal deterioration. *Neurology* **80,** 366–370 (2013).
72. Byrne, S., Heverin, M., Elamin, M., Walsh, C. & Hardiman, O. Intermediate repeat expansion length in C9orf72 may be pathological in amyotrophic lateral sclerosis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **15,** 148–150 (2014).
73. Rohrer, J. D. *et al.* C9orf72 expansions in frontotemporal dementia and amyotrophic lateral sclerosis. *Lancet Neurol* **14,** 291–301 (2015).
74. Gijselinck, I. *et al.* A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study. *Lancet Neurol* **11,** 54–65 (2012).
75. Millecamps, S. *et al.* Phenotype difference between ALS patients with expanded repeats in C9ORF72 and patients with mutations in other ALS-related genes. *J. Med. Genet.* **49,** 258–263 (2012).
76. Byrne, S. *et al.* Cognitive and clinical characteristics of patients with amyotrophic lateral sclerosis carrying a C9orf72 repeat expansion: a population-based cohort study. *Lancet Neurol* **11,** 232–240 (2012).
77. García-Redondo, A. *et al.* Analysis of the C9orf72 gene in patients with amyotrophic lateral sclerosis in Spain and different populations worldwide. *Hum. Mutat.* **34,** 79–82 (2013).
78. Sabatelli, M. *et al.* C9ORF72 hexanucleotide repeat expansions in the Italian sporadic ALS population. *Neurobiology of Aging* **33,** 1848.e15–20 (2012).
79. Simón-Sánchez, J. *et al.* The clinical and pathological phenotype of C9ORF72 hexanucleotide repeat expansions. *Brain* **135,** 723–735 (2012).
80. Mok, K. Y. *et al.* High frequency of the expanded C9ORF72 hexanucleotide repeat in familial and sporadic Greek ALS patients. *Neurobiology of Aging* **33,** 1851.e1–5 (2012).
81. Chester, C. *et al.* Rapidly progressive frontotemporal dementia and bulbar amyotrophic lateral sclerosis in Portuguese patients with C9orf72 mutation. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **14,** 70–72 (2013).
82. Boeve, B. F. *et al.* Characterization of frontotemporal dementia and/or amyotrophic lateral sclerosis associated with the GGGGCC repeat expansion in C9ORF72. *Brain* **135,** 765–783 (2012).
83. Takada, L. T. *et al.* Frontotemporal dementia in a Brazilian kindred with the c9orf72 mutation. *Arch Neurol* **69,** 1149–1153 (2012).
84. Dobson-Stone, C. *et al.* C9ORF72 repeat expansion in Australian and Spanish frontotemporal dementia patients. *PLoS ONE* **8,** e56899 (2013).
85. Cooper-Knock, J. *et al.* C9ORF72 transcription in a frontotemporal dementia case with two expanded alleles. *Neurology* **81,** 1719–1721 (2013).
86. Fratta, P. *et al.* Homozygosity for the C9orf72 GGGGCC repeat expansion in frontotemporal dementia. *Acta Neuropathol* **126,** 401–409 (2013).

87. Ishiura, H. *et al.* C9ORF72 repeat expansion in amyotrophic lateral sclerosis in the Kii peninsula of Japan. *Arch Neurol* **69,** 1154–1158 (2012).
88. Kim, E.-J. *et al.* Clinical and genetic analysis of MAPT, GRN, and C9orf72 genes in Korean patients with frontotemporal dementia. *Neurobiology of Aging* **35,** 1213.e13–7 (2014).
89. Jang, J.-H. *et al.* Analysis of the C9orf72 hexanucleotide repeat expansion in Korean patients with familial and sporadic amyotrophic lateral sclerosis. *Neurobiology of Aging* **34,** 1311.e7–9 (2013).
90. Zou, Z.-Y., Li, X.-G., Liu, M.-S. & Cui, L.-Y. Screening for C9orf72 repeat expansions in Chinese amyotrophic lateral sclerosis patients. *Neurobiology of Aging* **34,** 1710.e5–6 (2013).
91. Liu, R. *et al.* C9orf72 repeat expansions are not detected in Chinese patients with familial ALS. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* **14,** 630–631 (2013).
92. Ogaki, K. *et al.* Analysis of C9orf72 repeat expansion in 563 Japanese patients with amyotrophic lateral sclerosis. *Neurobiology of Aging* **33,** 2527.e11–6 (2012).
93. Ogaki, K. *et al.* Analyses of the MAPT, PGRN, and C9orf72 mutations in Japanese patients with FTLD, PSP, and CBS. *Parkinsonism Relat. Disord.* **19,** 15–20 (2013).
94. Konno, T. *et al.* Japanese amyotrophic lateral sclerosis patients with GGGGCC hexanucleotide repeat expansion in C9ORF72. *Journal of Neurology, Neurosurgery & Psychiatry* **84,** 398–401 (2013).
95. Tsai, C.-P. *et al.* A hexanucleotide repeat expansion in C9ORF72 causes familial and sporadic ALS in Taiwan. *Neurobiology of Aging* **33,** 2232.e11–2232.e18 (2012).
96. Benussi, L. *et al.* C9ORF72 hexanucleotide repeat number in frontotemporal lobar degeneration: a genotype-phenotype correlation study. *J. Alzheimers Dis.* **38,** 799–808 (2014).
97. Majounie, E. *et al.* Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol* **11,** 323–330 (2012).
98. Friedland, R. P. *et al.* Behavioral variant frontotemporal lobar degeneration with amyotrophic lateral sclerosis with a chromosome 9p21 hexanucleotide repeat. *Front Neurol* **3,** 136 (2012).
99. Akimoto, C. *et al.* A blinded international study on the reliability of genetic testing for GGGGCC-repeat expansions in C9orf72 reveals marked differences in results among 14 laboratories. *J. Med. Genet.* **51,** 419–424 (2014).
100. Harper, P. S., Harley, H. G., Reardon, W. & Shaw, D. J. *Anticipation in myotonic dystrophy: new light on an old problem. American journal of human genetics* **51,** 10–16 (1992).
101. Hübers, A. *et al.* Polymerase chain reaction and Southern blot-based analysis of the C9orf72 hexanucleotide repeat in different motor neuron diseases.

Neurobiology of Aging **35,** 1214.e1–6 (2014).
102. Chiò, A. *et al.* ALS/FTD phenotype in two Sardinian families carrying both C9ORF72 and TARDBP mutations. *Journal of Neurology, Neurosurgery & Psychiatry* **83,** 730–733 (2012).
103. Dols-Icardo, O. *et al.* Characterization of the repeat expansion size in C9orf72 in amyotrophic lateral sclerosis and frontotemporal dementia. *Hum. Mol. Genet.* **23,** 749–754 (2014).
104. van Blitterswijk, M. *et al.* Association between repeat sizes and clinical and pathological characteristics in carriers of C9ORF72 repeat expansions (Xpansize-72): a cross-sectional cohort study. *Lancet Neurol* **12,** 978–988 (2013).
105. López Castel, A., Cleary, J. D. & Pearson, C. E. Repeat instability as the basis for human diseases and as a potential target for therapy. *Nat. Rev. Mol. Cell Biol.* **11,** 165–170 (2010).
106. Nordin, A. *et al.* Extensive size variability of the GGGGCC expansion in C9orf72 in both neuronal and non-neuronal tissues in 18 patients with ALS or FTD. *Hum. Mol. Genet.* **24,** 3133–3142 (2015).
107. Harms, M. B. *et al.* Lack of C9ORF72 coding mutations supports a gain of function for repeat expansions in amyotrophic lateral sclerosis. *Neurobiology of Aging* **34,** 2234.e13–9 (2013).
108. Ratti, A. *et al.* C9ORF72 repeat expansion in a large Italian ALS cohort: evidence of a founder effect. *Neurobiology of Aging* **33,** 2528.e7–14 (2012).
109. Rollinson, S. *et al.* A small deletion in C9orf72 hides a proportion of expansion carriers in FTLD. *Neurobiology of Aging* (2014). doi:10.1016/j.neurobiolaging.2014.12.009
110. Koppers, M. *et al.* Screening for rare variants in the coding region of ALS-associated genes at 9p21.2 and 19p13.3. *Neurobiology of Aging* **34,** 1518.e5–7 (2013).
111. Suzuki, N. *et al.* The mouse C9ORF72 ortholog is enriched in neurons known to degenerate in ALS and FTD. *Nat. Neuroscience* **16,** 1725–1727 (2013).
112. Koppers, M. *et al.* C9orf72 ablation in mice does not cause motor neuron degeneration or motor deficits. *Ann Neurol.* **78,** 426–438 (2015).
113. Atkinson, R. A. K., Fernandez-Martos, C. M., Atkin, J. D., Vickers, J. C. & King, A. E. C9ORF72 expression and cellular localization over mouse development. *Acta Neuropathol Commun* **3,** 59 (2015).
114. Donnelly, C. J. *et al.* RNA Toxicity from the ALS/FTD C9ORF72 Expansion Is Mitigated by Antisense Intervention. *Neuron* **80,** 415–428 (2013).
115. Sareen, D., O'Rourke, J. G. & Meera, P. Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Science translational* (2013).
116. Waite, A. J. *et al.* Reduced C9orf72 protein levels in frontal cortex of amyotrophic lateral sclerosis and frontotemporal degeneration brain with the C9ORF72 hexanucleotide repeat expansion. *Neurobiology of Aging* **35,** 1779.e5–1779.e13 (2014).

117. Xiao, S. *et al.* Isoform-specific antibodies reveal distinct subcellular localizations of C9orf72 in amyotrophic lateral sclerosis. *Ann Neurol.* **78,** 568–583 (2015).
118. Levine, T. P., Daniels, R. D., Gatta, A. T., Wong, L. H. & Hayes, M. J. The product of C9orf72, a gene strongly implicated in neurodegeneration, is structurally related to DENN Rab-GEFs. *Bioinformatics* **29,** 499–503 (2013).
119. Zhang, D., Iyer, L. M., He, F. & Aravind, L. Discovery of Novel DENN Proteins: Implications for the Evolution of Eukaryotic Intracellular Membrane Structures and Human Disease. *Front Genet* **3,** 283 (2012).
120. Chaineau, M., Ioannou, M. S. & McPherson, P. S. Rab35: GEFs, GAPs and effectors. *Traffic* **14,** 1109–1117 (2013).
121. Farg, M. A. *et al.* C9ORF72, implicated in amytrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum. Mol. Genet.* **23,** 3579–3595 (2014).
122. Freibaum, B. D. *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525,** 129–133 (2015).
123. Jovičić, A. *et al.* Modifiers of C9orf72 dipeptide repeat toxicity connect nucleocytoplasmic transport defects to FTD/ALS. *Nat. Neuroscience* **18,** 1226–1229 (2015).
124. Zhang, K. *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525,** 56–61 (2015).
125. Woerner, A. C. *et al.* Cytoplasmic protein aggregates interfere with nucleo-cytoplasmic transport of protein and RNA. *Science* (2015).
126. Ivanov, P. *et al.* G-quadruplex structures contribute to the neuroprotective effects of angiogenin-induced tRNA fragments. *Proceedings of the National Academy of Sciences* **111,** 18201–18206 (2014).
127. Rossi, S. *et al.* Nuclear accumulation of mRNAs underlies G4C2-repeat-induced translational repression in a cellular model of C9orf72 ALS. *J. Cell. Sci.* **128,** 1787–1799 (2015).
128. Ling, S.-C., Polymenidou, M. & Cleveland, D. W. Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron* **79,** 416–438 (2013).
129. McMurray, C. T. Mechanisms of trinucleotide repeat instability during human development. *Nat. Rev. Genet.* **11,** 786–799 (2010).
130. La Spada, A. R. & Taylor, J. P. Repeat expansion disease: progress and puzzles in disease pathogenesis. *Nat. Rev. Genet.* **11,** 247–258 (2010).
131. Orr, H. T. & Zoghbi, H. Y. Trinucleotide Repeat Disorders. *Annu. Rev. Neurosci.* **30,** 575–621 (2007).
132. Mahadevan, M. *et al.* Myotonic dystrophy mutation: an unstable CTG repeat in the 3' untranslated region of the gene. *Science* **255,** 1253–1255 (1992).
133. Fu, Y. H. *et al.* An unstable triplet repeat in a gene related to myotonic muscular dystrophy. *Science* **255,** 1256–1258 (1992).
134. Brook, J. D. *et al.* Molecular basis of myotonic dystrophy: expansion of a trinucleotide (CTG) repeat at the 3' end of a transcript encoding a protein

kinase family member. *Cell* **69,** 385 (1992).
135. Kanadia, R. N. A Muscleblind Knockout Model for Myotonic Dystrophy. *Science* **302,** 1978–1980 (2003).
136. Orr, H. T. & Zoghbi, H. Y. Trinucleotide repeat disorders. *Annu. Rev. Neurosci.* **30,** 575–621 (2007).
137. Moseley, M. L. *et al.* Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat. Genet.* **38,** 758–769 (2006).
138. Michlewski, G. & Krzyzosiak, W. J. Molecular architecture of CAG repeats in human disease related transcripts. *J. Mol. Biol.* **340,** 665–679 (2004).
139. Costa, M. D. C. & Paulson, H. L. Toward understanding Machado-Joseph disease. *Progress in Neurobiology* **97,** 239–257 (2012).
140. Gaspar, C. *et al.* CAG tract of MJD-1 may be prone to frameshifts causing polyalanine accumulation. *Hum. Mol. Genet.* **9,** 1957–1966 (2000).
141. Toulouse, A. *et al.* Ribosomal frameshifting on MJD-1 transcripts with long CAG tracts. *Hum. Mol. Genet.* **14,** 2649–2660 (2005).
142. Stochmanski, S. J. *et al.* Expanded ATXN3 frameshifting events are toxic in Drosophila and mammalian neuron models. *Hum. Mol. Genet.* **21,** 2211–2218 (2012).
143. Ketteler, R. On programmed ribosomal frameshifting: the alternative proteomes. *Front Genet* **3,** 242 (2012).
144. Komar, A. A. & Hatzoglou, M. Cellular IRES-mediated translation. *Cell Cycle* **10,** 229–240 (2014).
145. Wojciechowska, M., Olejniczak, M., Galka-Marciniak, P., Jazurek, M. & Krzyzosiak, W. J. RAN translation and frameshifting as translational challenges at simple repeats of human neurodegenerative disorders. *Nucleic Acids Research* gku794 (2014).
146. A novel gene containing a trinucleotide repeat that is expanded and unstable on Huntington's disease chromosomes. The Huntington's Disease Collaborative Research Group. *Cell* **72,** 971–983 (1993).
147. Moseley, M. L. *et al.* Bidirectional expression of CUG and CAG expansion transcripts and intranuclear polyglutamine inclusions in spinocerebellar ataxia type 8. *Nat. Genet.* **38,** 758–769 (2006).
148. Zu, T. *et al.* Non-ATG-initiated translation directed by microsatellite expansions. *Proceedings of the National Academy of Sciences* **108,** 260–265 (2011).
149. Bañez-Coronel, M. *et al.* RAN Translation in Huntington Disease. *Neuron* **88,** 667–677 (2015).
150. Todd, P. K. *et al.* CGG Repeat-Associated Translation Mediates Neurodegeneration in Fragile X Tremor Ataxia Syndrome. *Neuron* 1–16 (2013).
151. Verkerk, A. J. *et al.* Identification of a gene (FMR-1) containing a CGG repeat coincident with a breakpoint cluster region exhibiting length variation in fragile X syndrome. *Cell* **65,** 905–914 (1991).

152. Hagerman, R. & Hagerman, P. Advances in clinical and molecular understanding of the FMR1 premutation and fragile X-associated tremor/ataxia syndrome. *The Lancet Neurology* **12,** 786–798 (2013).
153. Epstein, C. J. *et al.* Transgenic mice with increased Cu/Zn-superoxide dismutase activity: animal model of dosage effects in Down syndrome. *Proc. Natl. Acad. Sci. U.S.A.* **84,** 8044–8048 (1987).
154. Gurney, M. E. *et al.* Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation. *Science* **264,** 1772–1775 (1994).
155. Guo, Y.-S. *et al.* Sensory involvement in the SOD1-G93A mouse model of amyotrophic lateral sclerosis. *Exp Mol Med* **41,** 140–150 (2009).
156. Xu, Y.-F. *et al.* Wild-type human TDP-43 expression causes TDP-43 phosphorylation, mitochondrial aggregation, motor deficits, and early mortality in transgenic mice. *J. Neurosci.* **30,** 10851–10859 (2010).
157. Wils, H. *et al.* TDP-43 transgenic mice develop spastic paralysis and neuronal inclusions characteristic of ALS and frontotemporal lobar degeneration. *Proceedings of the National Academy of Sciences* **107,** 3858–3863 (2010).
158. Swarup, V. *et al.* Pathological hallmarks of amyotrophic lateral sclerosis/frontotemporal lobar degeneration in transgenic mice produced with TDP-43 genomic fragments. *Brain* **134,** 2610–2626 (2011).
159. Alfieri, J. A., Pino, N. S. & Igaz, L. M. Reversible behavioral phenotypes in a conditional mouse model of TDP-43 proteinopathies. *J. Neurosci.* **34,** 15244–15259 (2014).
160. Janssens, J. *et al.* Overexpression of ALS-associated p.M337V human TDP-43 in mice worsens disease features compared to wild-type human TDP-43 mice. *Mol Neurobiol* **48,** 22–35 (2013).
161. Wegorzewska, I., Bell, S., Cairns, N. J., Miller, T. M. & Baloh, R. H. TDP-43 mutant transgenic mice develop features of ALS and frontotemporal lobar degeneration. *Proceedings of the National Academy of Sciences* **106,** 18809–18814 (2009).
162. Mitchell, J. C. *et al.* Overexpression of human wild-type FUS causes progressive motor neuron degeneration in an age- and dose-dependent fashion. *Acta Neuropathol* **125,** 273–288 (2012).
163. Arnold, E. S. *et al.* ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. *Proceedings of the National Academy of Sciences* **110,** E736–45 (2013).
164. Verbeeck, C. *et al.* Expression of Fused in sarcoma mutations in mice recapitulates the neuropathology of FUS proteinopathies and provides insight into disease pathogenesis. *Molecular Neurodegeneration* **7,** 53 (2012).
165. Ittner, L. M. *et al.* FTD and ALS—translating mouse studies into clinical trials. *Nat Rev Neurol* **11,** 360–366 (2015).
166. Mitsumoto, H., Brooks, B. R. & Silani, V. Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? *Lancet Neurol* **13,** 1127–1138 (2014).

167. Poppe, L., Rué, L., Robberecht, W. & Van Den Bosch, L. Translating biological findings into new treatment strategies for amyotrophic lateral sclerosis (ALS). *Experimental Neurology* **262,** 138–151 (2014).
168. Su, Z. *et al.* Discovery of a Biomarker and Lead Small Molecules to Target r(GGGGCC)-Associated Defects in c9FTD/ALS. *Neuron* **83,** 1043–1050 (2014).
169. Wen, X. *et al.* Antisense Proline-Arginine RAN Dipeptides Linked to C9ORF72-ALS/FTD Form Toxic Nuclear Aggregates that Initiate In Vitro and In Vivo Neuronal Death. *Neuron* **84,** 1213–1225 (2014).
170. Kaletsky, R. & Murphy, C. T. The role of insulin/IGF-like signaling in C. elegans longevity and aging. *Dis Model Mech* **3,** 415–419 (2010).
171. Fay, D. Genetic mapping and manipulation: Chapter 1-Introduction and basics. *WormBook* (2006). doi:10.1895/wormbook.1.90.1
172. Hu, P. J. Dauer. *WormBook* (2007). doi:10.1895/wormbook.1.144.1
173. Lian Chew, Y., Fan, X., Götz, J. & Nicholas, H. R. Aging in the nervous system of Caenorhabditis elegans. *Commun Integr Biol* **6,** e25288 (2013).
174. C. elegans Sequencing Consortium. Genome sequence of the nematode C. elegans: a platform for investigating biology. *Science* **282,** 2012–2018 (1998).
175. Narayan, P., Ehsani, S. & Lindquist, S. Combating neurodegenerative disease with chemical probes and model systems. *Nat. Chem. Biol.* **10,** 911–920 (2014).
176. Murphy, C. T. & Hu, P. J. Insulin/insulin-like growth factor signaling in C. elegans. *WormBook* 1–43 (2013). doi:10.1895/wormbook.1.164.1
177. Gartner, A., Boag, P. R. & Blackwell, T. K. Germline survival and apoptosis. *WormBook* 1–20 (2008). doi:10.1895/wormbook.1.145.1
178. Nikoletopoulou, V. & Tavernarakis, N. Necrotic cell death in Caenorhabditis elegans. *Meth. Enzymol.* **545,** 127–155 (2014).
179. Ewbank, J. J. Signaling in the immune response. *WormBook* 1–12 (2006).
180. Altun, Z. F. & Hall, D. H. Nervous system, general description. *WormAtlas* (2011). doi:10.3908/wormatlas.1.18
181. Rand, J. B. Acetylcholine. *WormBook* 1–21 (2007).
182. Jorgensen, E. M. GABA. *WormBook* 1–13 (2005).
183. Rodriguez, M., Snoek, L. B., De Bono, M. & Kammenga, J. E. Worms under stress: C. elegans stress response and its relevance to complex human disease and aging. **29,** 367–374 (2013).
184. Lant, B. & Storey, K. B. An overview of stress response and hypometabolic strategies in Caenorhabditis elegans: conserved and contrasting signals with the mammalian system. *Int. J. Biol. Sci.* **6,** 9–50 (2010).
185. Yandell, M. D., Edgar, L. G. & Wood, W. B. Trimethylpsoralen induces small deletion mutations in Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U.S.A.* **91,** 1381–1385 (1994).
186. Brenner, S. The genetics of Caenorhabditis elegans. *Genetics* **77,** 71–94 (1974).
187. Zhuang, J. J. & Hunter, C. P. RNA interference in Caenorhabditis elegans:

uptake, mechanism, and regulation. *Parasitology* **139,** 560–573 (2012).
188. Calixto, A., Chelur, D., Topalidou, I., Chen, X. & Chalfie, M. Enhanced neuronal RNAi in C. elegans using SID-1. *Nat Meth* **7,** 554–559 (2010).
189. Firnhaber, C. & Hammarlund, M. Neuron-Specific Feeding RNAi in C. elegans and Its Use in a Screen for Essential Genes Required for GABA Neuron Function. *PLoS Genet* **9,** e1003921 (2013).
190. Qadota, H. *et al.* Establishment of a tissue-specific RNAi system in C. elegans. *Gene* **400,** 166–173 (2007).
191. Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. & Prasher, D. Green fluorescent protein as a marker for gene expression. *Science* **263,** 802–805 (1994).
192. Evans, T. C. Transformation and microinjection. *WormBook* 1–15 (2006).
193. Frøkjaer-Jensen, C. *et al.* Single-copy insertion of transgenes in Caenorhabditis elegans. *Nat. Genet.* **40,** 1375–1383 (2008).
194. Friedland, A. E. *et al.* Heritable genome editing in C. elegans via a CRISPR-Cas9 system. *Nat Meth* **10,** 741–743 (2013).
195. Vaccaro, A. *et al.* Mutant TDP-43 and FUS cause age-dependent paralysis and neurodegeneration in C. elegans. *PLoS ONE* **7,** e31321 (2012).
196. Murakami, T. *et al.* ALS mutations in FUS cause neuronal dysfunction and death in Caenorhabditis elegans by a dominant gain-of-function mechanism. *Hum. Mol. Genet.* **21,** 1–9 (2012).
197. Zhang, T., Mullane, P. C., Periz, G. & Wang, J. TDP-43 neurotoxicity and protein aggregation modulated by heat shock factor and insulin/IGF-1 signaling. *Hum. Mol. Genet.* **20,** 1952–1965 (2011).
198. Liachko, N. F., Guthrie, C. R. & Kraemer, B. C. Phosphorylation Promotes Neurotoxicity in a Caenorhabditis elegans Model of TDP-43 Proteinopathy. *Journal of Neuroscience* **30,** 16208–16219 (2010).
199. Ash, P. E. A. *et al.* Neurotoxic effects of TDP-43 overexpression in C. elegans. *Hum. Mol. Genet.* **19,** 3206–3218 (2010).
200. Oeda, T. *et al.* Oxidative stress causes abnormal accumulation of familial amyotrophic lateral sclerosis-related mutant SOD1 in transgenic Caenorhabditis elegans. *Hum. Mol. Genet.* **10,** 2013–2023 (2001).
201. Wang, J. *et al.* An ALS-Linked Mutant SOD1 Produces a Locomotor Defect Associated with Aggregation and Synaptic Dysfunction When Expressed in Neurons of Caenorhabditis elegans. *PLoS Genet* **5,** e1000350 (2009).
202. Gidalevitz, T., Krupinski, T., Garcia, S. & Morimoto, R. I. Destabilizing Protein Polymorphisms in the Genetic Background Direct Phenotypic Expression of Mutant SOD1 Toxicity. *PLoS Genet* **5,** e1000399 (2009).
203. Li, J. *et al.* Human superoxide dismutase 1 overexpression in motor neurons of Caenorhabditis elegans causes axon guidance defect and neurodegeneration. *Neurobiology of Aging* **35,** 837–846 (2014).
204. Vaccaro, A. *et al.* Pharmacological reduction of ER stress protects against TDP-43 neuronal toxicity in vivo. *Neurobiol. Dis.* **55,** 64–75 (2013).
205. Vaccaro, A. *et al.* TDP-1/TDP-43 regulates stress signaling and age-dependent proteotoxicity in Caenorhabditis elegans. *PLoS Genet* **8,** e1002806

(2012).
206. Zhang, T., Hwang, H. Y., Hao, H., Talbot, C. & Wang, J. Caenorhabditis elegans RNA-processing Protein TDP-1 Regulates Protein Homeostasis and Life Span. *Journal of Biological Chemistry* **287,** 8371–8382 (2012).
207. Zhang, T., Baldie, G., Periz, G. & Wang, J. RNA-Processing Protein TDP-43 Regulates FOXO-Dependent Protein Quality Control in Stress Response. *PLoS Genet* **10,** e1004693 (2014).
208. Tauffenberger, A., Chitramuthu, B. P., Bateman, A., Bennett, H. P. J. & Parker, J. A. Reduction of polyglutamine toxicity by TDP-43, FUS and progranulin in Huntington's disease models. *Hum. Mol. Genet.* **22,** 782–794 (2013).
209. Boyd, J. D. *et al.* A High-Content Screen Identifies Novel Compounds That Inhibit Stress-Induced TDP-43 Cellular Aggregation and Associated Cytotoxicity. *J Biomol Screen* (2013). doi:10.1177/1087057113501553
210. Du, M. & Wang, D. The neurotoxic effects of heavy metal exposure on GABAergic nervous system in nematode Caenorhabditis elegans. *Environ. Toxicol. Pharmacol.* **27,** 314–320 (2009).
211. Negga, R. *et al.* Exposure to glyphosate- and/or Mn/Zn-ethylene-bis-dithiocarbamate-containing pesticides leads to degeneration of γ-aminobutyric acid and dopamine neurons in Caenorhabditis elegans. *Neurotox Res* **21,** 281–290 (2012).
212. Estevez, A. O., Morgan, K. L., Szewczyk, N. J., Gems, D. & Estevez, M. The neurodegenerative effects of selenium are inhibited by FOXO and PINK1/PTEN regulation of insulin/insulin-like growth factor signaling in Caenorhabditis elegans. *Neurotoxicology* (2014).
213. Alexander, A. G., Marfil, V. & Li, C. Use of Caenorhabditis elegans as a model to study Alzheimer's disease and other neurodegenerative diseases. *Front Genet* **5,** 279 (2014).
214. Harrington, A. J., Hamamichi, S., Caldwell, G. A. & Caldwell, K. A. C. elegans as a model organism to investigate molecular pathways involved with Parkinson's disease. *Dev. Dyn.* **239,** 1282–1295 (2010).
215. Brignull, H. R., Morley, J. F., Garcia, S. M. & Morimoto, R. I. Modeling polyglutamine pathogenesis in C. elegans. *Meth. Enzymol.* **412,** 256–282 (2006).
216. Blackstone, C. Cellular Pathways of Hereditary Spastic Paraplegia *. *Annu. Rev. Neurosci.* **35,** 25–47 (2012).
217. Dion, P. A., Daoud, H. & Rouleau, G. A. Genetics of motor neuron disorders: new insights into pathogenic mechanisms. *Nat. Rev. Genet.* **10,** 769–782 (2009).
218. Matsushita-Ishiodori, Y., Yamanaka, K. & Ogura, T. The C. elegans homologue of the spastic paraplegia protein, spastin, disassembles microtubules. *Biochemical and Biophysical Research Communications* **359,** 157–162 (2007).
219. Zhao, J. *et al.* Hereditary spastic paraplegia-associated mutations in the

NIPA1 gene and its Caenorhabditis elegans homolog trigger neural degeneration in vitro and in vivo through a gain-of-function mechanism. *J. Neurosci.* **28,** 13938–13951 (2008).
220. Arnold, W. D. & Burghes, A. H. M. Spinal muscular atrophy: development and implementation of potential treatments. *Ann Neurol.* **74,** 348–362 (2013).
221. Miguel-Aliaga, I. *et al.* The Caenorhabditis elegans orthologue of the human gene responsible for spinal muscular atrophy is a maternal product critical for germline maturation and embryonic viability. *Hum. Mol. Genet.* **8,** 2133–2143 (1999).
222. Kwon, D. Y. *et al.* The E3 ubiquitin ligase mind bomb 1 ubiquitinates and promotes the degradation of survival of motor neuron protein. *Mol. Biol. Cell* **24,** 1863–1871 (2013).
223. Chen, K.-Y. *et al.* Length-dependent toxicity of untranslated CUG repeats on Caenorhabditis elegans. *Biochemical and Biophysical Research Communications* **352,** 774–779 (2007).
224. Wang, L.-C. *et al.* Muscleblind participates in RNA toxicity of expanded CAG and CUG repeats in Caenorhabditis elegans. *Cell. Mol. Life Sci.* **68,** 1255–1267 (2011).
225. Spilker, K. A., Wang, G. J., Tugizova, M. S. & Shen, K. Caenorhabditis elegans Muscleblind homolog mbl-1 functions in neurons to regulate synapse formation. *Neural Dev* **7,** 7 (2012).
226. Sareen, D. *et al.* Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci Transl Med* **5,** 208ra149–208ra149 (2013).
227. Belzil, V. V. *et al.* Characterization of DNA hypermethylation in the cerebellum of c9FTD/ALS patients. *Brain Research* **1584,** 15–21 (2014).
228. Russ, J. *et al.* Hypermethylation of repeat expanded C9orf72 is a clinical and molecular disease modifier. *Acta Neuropathol* **129,** 39–52 (2015).
229. Xi, Z. *et al.* Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. *Am. J. Hum. Genet.* **92,** 981–989 (2013).
230. Xi, Z. *et al.* Hypermethylation of the CpG-island near the C9orf72 G_4C_2-repeat expansion in FTLD patients. *Hum. Mol. Genet.* **23,** 5630–5637 (2014).
231. Xi, Z. *et al.* The C9orf72 repeat expansion itself is methylated in ALS and FTLD patients. *Acta Neuropathol* **129,** 715–727 (2015).
232. Therrien, M., Rouleau, G. A., Dion, P. A. & Parker, J. A. Deletion of C9ORF72 Results in Motor Neuron Degeneration and Stress Sensitivity in C. elegans. *PLoS ONE* **8,** e83450 (2013).
233. Reddy, K., Zamiri, B., Stanley, S. Y. R., Macgregor, R. B. & Pearson, C. E. The disease-associated r(GGGGCC)n repeat from the C9orf72 gene forms tract length-dependent uni- and multimolecular RNA G-quadruplex structures. *Journal of Biological Chemistry* **288,** 9860–9866 (2013).
234. Reddy, K. *et al.* Processing of double-R-loops in (CAG)·(CTG) and C9orf72 (GGGGCC)·(GGCCCC) repeats causes instability. *Nucleic Acids Research* gku658 (2014).

235. Mori, K. *et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339,** 1335–1338 (2013).
236. Ash, P. E. A. *et al.* Unconventional translation of C9ORF72 GGGGCC expansion generates insoluble polypeptides specific to c9FTD/ALS. *Neuron* **77,** 639–646 (2013).
237. Polymenidou, M. *et al.* Misregulated RNA processing in amyotrophic lateral sclerosis. *Brain Research* **1462,** 3–15 (2012).
238. Lagier-Tourenne, C. *et al.* Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proceedings of the National Academy of Sciences* **110,** E4530–9 (2013).
239. O'Reilly, L. P., Luke, C. J., Perlmutter, D. H., Silverman, G. A. & Pak, S. C. C. elegans in high-throughput drug discovery. *Adv. Drug Deliv. Rev.* (2013).
240. Boucher, B. & Jenna, S. Genetic interaction networks: better understand to better predict. *Front Genet* **4,** 290 (2013).
241. Xu, Z. *et al.* Expanded GGGGCC repeat RNA associated with amyotrophic lateral sclerosis and frontotemporal dementia causes neurodegeneration. *Proceedings of the National Academy of Sciences* **110,** 7778–7783 (2013).
242. Mizielinska, S. *et al.* C9orf72 repeat expansions cause neurodegeneration in Drosophila through arginine-rich proteins. *Science* 1256800 (2014).
243. Tran, H. *et al.* Differential Toxicity of Nuclear RNA Foci versus Dipeptide Repeat Proteins in a Drosophila Model of C9ORF72 FTD/ALS. *Neuron* **87,** 1207–1214 (2015).
244. Burguete, A. S. *et al.* GGGGCC microsatellite RNA is neuritically localized, induces branching defects, and perturbs transport granule function. *eLife* **4,** 1–58 (2015).
245. Chew, J. *et al.* Neurodegeneration. C9ORF72 repeat expansions in mice cause TDP-43 pathology, neuronal loss, and behavioral deficits. *Science* **348,** 1151–1154 (2015).
246. Peters, O. M. *et al.* Human C9ORF72 Hexanucleotide Expansion Reproduces RNA Foci and Dipeptide Repeat Proteins but Not Neurodegeneration in BAC Transgenic Mice. *Neuron* **88,** 902–909 (2015).
247. O'Rourke, J. G. *et al.* C9orf72 BAC Transgenic Mice Display Typical Pathologic Features of ALS/FTD. *Neuron* **88,** 892–901 (2015).
248. Moronetti Mazzeo, L. E., Dersh, D., Boccitto, M., Kalb, R. G. & Lamitina, T. Stress and aging induce distinct polyQ protein aggregation states. *Proceedings of the National Academy of Sciences* **109,** 10587–10592 (2012).
249. Verbalis, J. G. How Does the Brain Sense Osmolality? *Journal of the American Society of Nephrology* **18,** 3056–3059 (2007).
250. Cooper, J. K. *et al.* Truncated N-terminal fragments of huntingtin with expanded glutamine repeats form nuclear and cytoplasmic aggregates in cell culture. *Hum. Mol. Genet.* **7,** 783–790 (1998).
251. Haacke, A. *et al.* Proteolytic cleavage of polyglutamine-expanded ataxin-3 is critical for aggregation and sequestration of non-expanded ataxin-3. *Hum. Mol. Genet.* **15,** 555–568 (2006).

252. Pouladi, M. A., Morton, A. J. & Hayden, M. R. Choosing an animal model for the study of Huntington's disease. *Nat Rev Neurosci* **14,** 708–721 (2013).
253. Couthouis, J. *et al.* A yeast functional screen predicts new candidate ALS disease genes. *Proceedings of the National Academy of Sciences* **108,** 20881–20890 (2011).
254. Couthouis, J. *et al.* Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **21,** 2899–2911 (2012).
255. Kino, Y. *et al.* FUS/TLS deficiency causes behavioral and pathological abnormalities distinct from amyotrophic lateral sclerosis. *Acta Neuropathol Commun* **3,** 24 (2015).
256. Ciura, S. *et al.* Loss of function of C9orf72 causes motor deficits in a zebrafish model of Amyotrophic Lateral Sclerosis. *Ann Neurol.* 180–187 (2013).
257. Liu, E. Y. *et al.* C9orf72 hypermethylation protects against repeat expansion-associated pathology in ALS/FTD. *Acta Neuropathol* (2014).
258. McMillan, C. T. *et al.* C9orf72 promoter hypermethylation is neuroprotective: Neuroimaging and neuropathologic evidence. *Neurology* 10.1212/WNL.0000000000001495 (2015).
259. Gendron, T. F. *et al.* Antisense transcripts of the expanded C9ORF72 hexanucleotide repeat form nuclear RNA foci and undergo repeat-associated non-ATG translation in c9FTD/ALS. *Acta Neuropathol* **126,** 829–844 (2013).
260. Mori, K. *et al.* Bidirectional transcripts of the expanded C9orf72 hexanucleotide repeat are translated into aggregating dipeptide repeat proteins. *Acta Neuropathol* **126,** 881–893 (2013).
261. Su, Z. *et al.* Discovery of a Biomarker and Lead Small Molecules to Target r(GGGGCC)-Associated Defects in c9FTD/ALS. *Neuron* **83,** 1043–1050 (2014).
262. Kwon, I. *et al.* Poly-dipeptides encoded by the C9orf72 repeats bind nucleoli, impede RNA biogenesis, and kill cells. *Science* **345,** 1139–1145 (2014).
263. Mackenzie, I. R. *et al.* Dipeptide repeat protein pathology in C9ORF72 mutation cases: clinico-pathological correlations. *Acta Neuropathol* **126,** 859–879 (2013).
264. Cooper-Knock, J. *et al.* Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions. *Brain* (2014).
265. Lee, Y.-B. *et al.* Hexanucleotide Repeats in ALS/FTD Form Length-Dependent RNA Foci, Sequester RNA Binding Proteins, and Are Neurotoxic. *CellReports* **5,** 1178–1186 (2013).
266. Hokkanen, S. *et al.* Lack of Pur-alpha alters postnatal brain development and causes megalencephaly. *Hum. Mol. Genet.* **21,** 473–484 (2012).
267. Mann, D. M. A. *et al.* Dipeptide repeat proteins are present in the p62 positive inclusions in patients with frontotemporal lobar degeneration and motor neurone disease associated with expansions in C9ORF72. *Acta Neuropathol Commun* **1,** 68 (2013).
268. Zhang, S., Feany, M. B., Saraswati, S., Littleton, J. T. & Perrimon, N. Inactivation of Drosophila Huntingtin affects long-term adult functioning and

the pathogenesis of a Huntington's disease model. *Dis Model Mech* **2,** 247–266 (2009).
269. Di Prospero, N. A. & Tagle, D. A. Normal and mutant huntingtin: partners in crime. *Nature Medicine* **6,** 1208–1209 (2000).
270. Yang, H. *et al.* Aggregation of polyglutamine-expanded ataxin-3 sequesters its specific interacting partners into inclusions: implication in a loss-of-function pathology. *Sci. Rep.* **4,** 6410 (2014).
271. Bensimon, G., Lacomblez, L. & Meininger, V. A controlled trial of riluzole in amyotrophic lateral sclerosis. ALS/Riluzole Study Group. *N Engl J Med* **330,** 585–591 (1994).
272. Haddad, H. *et al.* Riluzole attenuates spinal muscular atrophy disease progression in a mouse model. *Muscle Nerve* **28,** 432–437 (2003).
273. Genç, B. & Ozdinler, P. H. Moving forward in clinical trials for ALS: motor neurons lead the way please. *Drug Discov. Today* **19,** 441–449 (2014).
274. Sreedharan, J. & Brown, R. H. Amyotrophic lateral sclerosis: Problems and prospects. *Ann Neurol.* **74,** 309–316 (2013).
275. Swinney, D. C. & Anthony, J. How were new medicines discovered? *Nat Rev Drug Discov* **10,** 507–519 (2011).
276. Vincent, F. *et al.* Developing predictive assays: the phenotypic screening "rule of 3". *Sci Transl Med* **7,** 293ps15–293ps15 (2015).
277. Long, C. *et al.* Postnatal genome editing partially restores dystrophin expression in a mouse model of muscular dystrophy. *Science* **351,** 400–403 (2016).
278. Nelson, C. E. *et al.* In vivo genome editing improves muscle function in a mouse model of Duchenne muscular dystrophy. *Science* **351,** 403–407 (2016).
279. Tabebordbar, M. *et al.* In vivo gene editing in dystrophic mouse muscle and muscle stem cells. *Science* **351,** 407–411 (2016).

ANNEXE I

MANUSCRIPT: *FET PROTEINS REGULATE LIFESPAN AND NEURONAL INTEGRITY*

AUTHORS AND AFFILIATIONS:

Martine Therrien[1,2], Guy A. Rouleau[3,4], Patrick A. Dion[3,4] and J. Alex Parker[1,5]*

Affiliations:
[1] CHUM Research Center, Montreal, H2X 3H8, Canada
[2] Pathology and Cell biology department, University of Montreal, Montreal, H3T 1J4, Canada.
[3] Neurology and Neurosurgery department, McGill University, Montreal, H3A 0G4, Canada.
[4] Montreal Neurological Hospital, Montreal, H3A 2B4, Canada.
[5] Department of Neuroscience, University of Montreal, Montreal, H3T 1J4, Canada.
***Corresponding Author:** JA.P

AUTHOR CONTRIBUTIONS

MT and JAP conceived the study

MT performed the experiments

MT and JAP wrote the manuscript

MT, JAP, GAR and PAD reviewed the manuscript.

** manuscript under review at Scientific Reports

ABSTRACT:

The FET protein family includes FUS, EWS and TAF15 proteins, all of which have been linked to amyotrophic lateral sclerosis, a fatal neurodegenerative disease affecting motor neurons. Here, we show that a reduction of FET proteins in the nematode *Caenorhabditis elegans* causes synaptic dysfunction accompanied by impaired motor phenotypes. FET proteins are also involved in the regulation of lifespan and stress resistance, acting partially through the insulin/IGF-signalling pathway. We propose that FET proteins are essential for the maintenance of lifespan, cellular stress resistance and neuronal integrity.

INTRODUCTION

Mutations in fused-in-sarcoma (*FUS*) are one of the causes of amyotrophic lateral sclerosis (ALS) (1-3), a fatal neurodegenerative disease causing loss of motor neurons. Mutations are found in different domains of the protein and cause its cellular mislocalization (3). The presence of FUS in the cytoplasm suggests a gain of toxic function mechanism, but the depletion of FUS from the nucleus also points to a loss of normal function being implicated in motor neuron degeneration. Due to its nature as an RNA/DNA binding protein, FUS has been shown to participate in many cellular functions including translation, splicing, and RNA transport (4). FUS is part of the FET protein family that includes two other RNA binding proteins; Ewin sarcoma breakpoint region 1 (*EWSR1* gene encoding the EWS protein) and TBP associated factor 15 (*TAF15*). These proteins are highly similar and it is thought that they share common functions(5, 6). Furthermore, *EWSR1* and *TAF15* mutations have been linked to some sporadic cases of ALS(7, 8). However, how mutant FET proteins cause neuronal loss is still unclear.

Many proteins associated with ALS are evolutionarily conserved in the nematode *Caenorhabditis elegans. C. elegans* are transparent nematodes that have been used to make important contributions to the fields of neuroscience and aging. More recently *C. elegans* has emerged as a useful model to study human diseases, namely conserved aspects of age-dependent neurodegeneration(9).

To better understand the function of FET proteins, we characterised the *fust-1(tm4439)* deletion mutant in worms. *C. elegans* has a simple, largely non-redundant genome and many highly conserved human genes have a single orthologue in the nematode. Here,

fust-1 is the orthologue of *FUS*, *EWSR1* and *TAF15*. Using a loss of function mutation, we show that *fust-1* is a key gene acting to regulate neuronal integrity, lifespan and cellular stress responses. Also, for some of these functions, *fust-1* is an active component of the insulin/IGF-like signalling pathway (ISS).

RESULTS

FUST-1 is required for neuronal integrity. To understand the function and role of the FET proteins, we characterised *fust-1*, which encodes the *C. elegans* orthologue of *FUS, EWSR1, TAF15.* At the protein level, FUST-1 shares 50% identity with the FUS human protein, 32% identity with EWS and 35% identity with TAF15. Bioinformatic analyses using NetNes (http://www.cbs.dtu.dk/services/NetNES/) and Prosite (http://prosite.expasy.org/scanprosite/) confirmed the conservation the main functional domains of FET proteins including the RNA binding domain, the zinc finger motif and the nuclear export signal (Fig. 1A and Supplementary Material, Fig. S1A and B). To investigate the role of FUST-1, we used a *C. elegans* deletion mutant strain, *fust-1(tm4439)* which contains 240 base pair deletion. The deletion mutant worms *fust-1(tm4439)* show 50% decreased expression of *fust-1* compared wild-type N2 worms (Supplementary Material, Fig. S1C), suggesting that those worms could be used to model the effect of a loss of function of FUST-1.

Previous reports studying the function of the ortholog of *FUS* in *Drosophila, Cabeza,* have suggested that a decreased expression of *Cabeza* in flies induces loss of neurons(10-13). To evaluate if this function was conserved in *C. elegans*, deletion mutant worms were evaluated for age-dependent paralysis, a motor phenotype that has been shown to be a good predictor of neuronal integrity(14-16). At day 1 of adulthood, *fust-1(tm4439*) mutants had a normal motor behaviour when compared to wild-type N2 worms, but as the mutants aged they showed progressive motility defects leading to paralysis, reaching

66% paralysis by day 12 of adulthood compare to the 13% observed for wild-type N2 controls (Fig. 1B and Supplementary Material, Table S1A). We have previously shown that the paralysis phenotype that occurs when worms are grown on solid media after several days can be observed within hours when the worms are grown in liquid culture(17, 18). The swimming behaviour of *C. elegans* is an energetically costly activity that actively engages the neuromuscular junction and may be a phenotype relevant to the study of the health of motor neurons. To study the movement of the animals in liquid, worms were placed in a 96-well plate and their movements were evaluated with an automated method that measures locomotion activity based on infrared beam scattering(19). With this method, *fust-1* mutants initially exhibit normal motility behaviour but show a drastic decrease in movement over time (Fig. 1C and Supplementary Material, Table S1B).

Next we assessed the integrity of GABAergic motor neurons using an *unc-47p*::mCherry reporter strain(20). UNC-47 is a GABA vesicular transporter and is expressed in the 26 GABAergic motor neurons of the worm (21). At day 9, *fust-1(tm4439)* mutants exhibit an increase in the number of gaps or breaks along the ventral cord (Fig. 1D and E) which coincides with the onset of the paralysis phenotype.

To evaluate potential synaptic dysfunction, an *unc25p::snb-1::GFP* reporter strain was used. SNB-1 is a synaptic vesicle protein and has been used to visualized synapses(22). Starting at day 1 of adulthood, *fust-1* deletion mutant worms exhibit abnormal organization of SNB-1 protein compared to wild-type worms. SNB-1 abnormal localization affected motor neurons that also exhibit gaps along their axons (Fig. 1D ii) or

could affect neuron prior to breakage of the axons (Fig. 1D iii). The proportion of worms with abnormal SNB-1 localization increased with aging reaching 60% of the worms at day 9 compared to 35% of the wild-type worms (Fig. 1F).

To evaluate the health of the neuromuscular junction, worms were exposed to aldicarb (2-methyl-2(methylthio) propanal o-[(methylamino)-carbonyl] oxime), an acetylcholine esterase inhibitor that causes the build-up of acetyl choline at the neuromuscular junction leading to paralysis(23). Worms with defects in vesicular release at the neuromuscular junction exhibit hypersensitivity, similar to *unc-47* mutants(24), or hyper-resistance, similar to *unc-64/Syntaxin* mutants(25). *fust-1* mutants exhibit hypersensitivity on aldicarb when compare to wild-type N2, reaching 80% paralysis compared to 40% for the wild-type control after two hours in aldicarb (Fig. 1G). These data suggest an abnormal function of the neuromuscular junction perhaps due to a decrease of GABA release in the *fust-1* mutants.

To confirm that the effects observed in our *fust-1(tm4439)* worms were due to the loss of function of *fust-1*, we generated a transgenic worm expressing full-length *fust-1* linked to GFP under the control of its own promoter (*fust-1p::fust-1::GFP*). This strain exhibits GFP expression throughout development and adulthood with expression in the head, pharynx, intestine and tail of the adult worm (Supplementary Material, Fig. S1H). When crossed to the mutant worms, the *fust-1p::fust-1::GFP* construct completely rescued the paralysis phenotype of the *fust-1(tm4439*) worms (Fig. 1H and Supplementary Material, Table S1C) suggesting that indeed the motor phenotype is due to a loss of *fust-1*.

Overall, these results suggest that synaptic dysfunction precedes neuronal loss and that aging could promote the development of the motor phenotype observed in the *fust-1(tm4439*) deletion mutant worms.

FUST-1 is involved in lifespan regulation. Genetic signalling pathways regulating aging have been extensively studied in *C. elegans* and central to lifespan and stress response mechanisms is the insulin /IGF-like signalling pathway (IIS)(26). DAF-2 is the sole insulin/IGF receptor in *C. elegans* and hypomorphic *daf-2* mutants are long-lived and highly resistant to environmental stress(26). We constructed a *daf-2(e1370); fust-1(tm4439)* double mutant strain and observed that the loss of *fust-1* completely abolished the extended lifespan phenotype of *daf-2(e1370)* mutants (Fig. 2A and Supplementary Material, Table S2A). These data suggest that that *fust-1* functions within the IIS to regulate longevity.

In *C. elegans*, a crucial downstream effector of the IIS is the forkhead box O (FOXO) transcription factor encoded by *daf-16*. The long-lived phenotypes of *daf-2* mutants is completely dependent on *daf-16*(26). The *daf-16(mu86)* mutants are short-lived but *fust-1(tm4439)*; *daf-16(mu86)* double mutants have lifespan similar to *daf-16(mu86)* mutants alone (Fig. 2B and Supplementary Material, Table S2B). These data suggest that regulation of lifespan by *fust-1* is dependent on *daf-16* and the decreased expression of *fust-1* does not affect the lifespan of *daf-16(mu86)*.

We observed that although *fust-1(tm4439*) mutants have a normal lifespan at 20°C and 25°C (Supplementary Material, Fig. S3A and B and Table S2C and D), the overexpression of *fust-1* caused an increased lifespan compared to wild-type worms (Fig. 2C, Supplementary Material Table S2E). Additionally, the overexpression of *fust-1* had an additive effect on *daf-2* mutant lifespan (Fig. 2D, Supplementary Material Table S2F). Thus our data suggest that *fust-1* is essential regulating lifespan via the IIS and that lifespan-extension is modulated by FUST-1 in a dose-dependent manner.

FUST-1 is involved in resistance to environmental stress. Another important role of the IIS is the regulation of cellular stress response. Worms were tested against different environmental stresses to evaluate the contribution of *fust-1* within the IIS. First, worms were exposed to juglone (5-hydroxy-1,4-naphthoquinone), a natural product that causes the production of intracellular free radical in worms causing an acute oxidative stress(27). The *fust-1* deletion mutants were more sensitive than wild-type N2 worms (Fig. 3A) and the sensitivity was rescued by the *fust-1p::fust-1::GFP* transgene in *fust-1(tm4439*) mutants (Supplementary Material, Fig. 4A). Next we examined *daf-2(e1370); fust-1(tm4439)* double mutants and observed that these animals were more sensitive to oxidative stress than *daf-2(e1370)* mutants, but more resistant than *fust-1* mutants alone (Fig. 3B). These data suggest that the IIS pathway is only partially reliant on *fust-1* in response to oxidative stress. These data suggest that *fust-1* is involved in the IIS.

fust-1 deletion mutants were then tested for their resistance to osmotic stress. *fust-1* mutants showed sensitivity to a hypertonic environment induced by NaCl (Fig 3C) and

this phenotype was partially rescued by the overexpression of *fust-1* (Supplementary Material, Fig S4B). The deletion of *fust-1* did not affect the sensitivity of *daf-2* mutants to osmotic stress (Fig 3D) suggesting that *fust-1* is completely independent of the IIS for the regulation of osmotic stress.

Finally, worms were exposed to thermal stress. The *fust-1* deletion mutants had a normal sensitivity when submitted to 37°C even after 14 hours (Supplementary Material, Fig. S4C). Also, the deletion of *fust-1* in *daf-2* mutants did not significantly change their resistance to this stress (Supplementary Material, Fig. S4D) suggesting that *fust-1* does not participate in responses to thermal stress.

***fust-1* expression is regulated by the IIS.** Genes participating in the IIS are known to have their expressions modulated under stress conditions or in IIS mutants(28),(29). To test if environmental stresses could induce the expression of *fust-1*, a transcriptional reporter of *fust-1* (*fust-1p::GFP*) (Supplementary Material, Fig. S5) was used to specifically evaluate the gene expression profile of *fust-1* and not the protein stability or degradation under these conditions. Worms were submitted to oxidative and osmotic stresses, the two types of stresses where *fust-1* seems to be the most involved. Osmotic but not oxidative could induce the expression of *fust-1* (Fig. 4A and B). To evaluate if *fust-1* expression is regulated by the IIS, expression level was measured in IIS mutant worms. When measured by qRT-PCR the *daf-2* mutants exhibit a two-fold higher expression level of *fust-1* than wild-type N2 (Fig. 4C). However, *daf-16* mutants exhibit a decreased expression level of *fust-1* (Fig. 4C) suggesting that *fust-1* is overexpressed when the IIS is

reduced. Therefore, IIS pathway mutants have an abnormal expression of *fust-1* and osmotic stress can induce *fust-1* expression independently.

***fust-1* involvement in neuronal integrity is independent of the IIS.** The IIS is involved in maintaining neuronal integrity(30) and some reports have previously suggested a link between the IIS and neurodegeneration in *C. elegans*(31, 32) . Therefore the impact of the *daf-16* mutants was tested on *fust-1* motor phenotypes. The lack of *daf-16* did not increase nor decreased the paralysis of the *fust-1(tm4439*) mutants (Fig. 5A). In previous studies, it was shown that stress sensitivity could cause neurodegeneration(14). In order to test this hypothesis, worms were tested against oxidative, osmotic or thermal stresses and neurodegeneration of the GABAergic neurons was evaluated. Oxidative and osmotic stresses cause neurodegeneration in the *fust-1* deletion mutants but thermal stress had no significant effect (Fig. 5B), suggesting a link between the role of *fust-1* in neuronal integrity and the survival to these stresses. However, the lack *daf-16* did not change the percentage of animal with neurodegeneration (Fig. 5B). Therefore, *fust-1* seems to be involved in neuronal integrity independently of the IIS.

DISCUSSION

FUS, EWS and TAF15 are RNA binding proteins that form the FET protein family. FET proteins were first identified as being involved in tumorigenensis, causing DNA translocation (6). Mutations in FET proteins have recently been linked to ALS (3, 7, 8). In ALS, most mutations affecting these genes are missense mutations and cause their mislocalization from the nucleus to the cytoplasm(1, 2, 7, 8) The effect of this mislocalization and how the mutant proteins cause toxicity to motor neurons is unknown. Because of their high sequence and domain similarity, all three proteins have the same orthologue in the nematode *C. elegans*, encoded by the gene *fust-1*. Using a deletion mutant strain, we have shown here that a decrease in *fust-1* expression can cause impaired stress resistance, lifespan regulation and neuronal integrity.

Using *C. elegans*, other ALS proteins have been reported to be involved in stress and lifespan regulation. *tdp-1*, the orthologue of *TDP-43*, and *alfa-1*, the orthologue of *C9orf72,* have been shown to impair stress response and TDP-1 was shown the be a key regulator of longevity(17, 33-35). In human, TDP-43 localization was shown to change upon stress where wild-type TDP-43 relocalizes in stress granules(36, 37). In humans, FUS was also shown to participate to stress granules formation and exhibit a change in localization upon stress induction(38, 39). More specifically, FUS localization changes after osmotic stress and a reduction of expression of *FUS* in human cell lines induced a loss of cell viability in a hyperosmolar environment(40). EWS and TAF-15 were also shown to translocate to stress granules in human cells and the RNA binding domain of

these proteins seem essential for this phenomenon(41, 42). Together, these results suggest the FET proteins are involved in stress response in *C. elegans* and humans.

The *Drosophila* gene *Cabeza* is the orthologue of *FUS*, *TAF15* and *EWSR1* (from http://flybase.org/reports/FBgn0011571.html). Interestingly in *Drosophila*, *Cabeza* was also shown to be involved in lifespan, neuronal integrity, synaptic function(10, 11, 13, 43). Synaptic dysfunctions were also shown to precede neuronal loss recapitulating features observed in *C. elegans*. In a recent *FUS* knock-out model, the mice with a complete loss of expression of *FUS* exhibit change in behaviour but no motor neuron loss(44). However, whole transcriptome analysis of spinal cord tissue of these mice has shown an increase in *EWSR1* and *TAF15* expression(44). It is known that FUS can regulate itself(45) and it seems that it can also regulate expression of protein with similar function such as TAF-15 and EWS-1. Therefore, using simple model organisms where compensatory mechanisms are less frequent we can reveal functions of protein family more easily and suggest here that the FET proteins act together to maintain neuron integrity and lifespan.

Using the *fust-1* deletion mutant, we have also shown that the FET protein family is involved in the IIS to maintain longevity and stress resistance. Also, the FUST-1 expression level was dysregulated in IIS mutants suggesting that a reduction in the pathway causes an overexpression of FUST-1. In human, IIS exerts its effect through many effectors including its many receptors, the kinases PI3K and AKT as well as FOXO transcription factors. TAF-15, EWS and FUS all contain a RNA binding domain suggestive of a role in RNA expression and metabolism. In a previous study, using UV cross-linking

immunoprecipitation followed by whole transcriptome sequencing (CLIP-seq), many RNA targets of the FET protein family have been revealed. Interestingly, members of the AKT and FOXO protein family were identified as targets of wild type TAF-15, EWS and FUS proteins in human cells(5). Using different *C. elegans* mutants, we have shown genetic interactions between the FET proteins and members of the IIS, results in human further our hypothesis and suggest direct interactions with members of the IIS.

In conclusion, using *C. elegans* deletion mutants, several functions of the FET proteins have been revealed. FUST-1 is actively participating and regulated by the IIS for the maintenance of lifespan and for proper resistance to oxidative stress. Independently of the IIS, FUST-1 is involved in neuronal integrity and the osmotic stress response suggesting its participation in other stress response pathways. It is still unclear if mutations of the FET protein family causative of ALS cause a loss of function or a gain of function. However, mutations in essential tremor, another neurodegenerative disorder, were found in *FUS* and are suggestive of a loss of function mechanism (46). Therefore, therapies targeting one of those proteins will have to be highly specific and take into account the impact on the other members of the FET family.

MATERIALS AND METHODS

Strains and maintenance

Standard method for culturing and handling the worms were used(47). Worms were cultured on standard NGM media streak with OP-50 *Escherichia. coli* strain at 20°C if not specified otherwise. For a list of strains used see Supplementary Table 4. *fust-1p*::*fust*-1::GFP fosmid was obtained by the TransgenOme project (48) and confirmed by sequencing. Strain was created by microinjection in *unc-119(ed3)* worms.

Bioinformatic analyses

For amino acid alignment, FUST-1 protein sequence was used as query sequence (Wormbase) and compared to FUS (CCDS58454.1), EWS isoform (CCDS54513.1) and TAF15 (CCDS59279.1) as subject sequence and align using BlastP (http://blast.ncbi.nlm.nih.gov/blast/Blast.cgi?PROGRAM=blastp&PAGE_TYPE=BlastSearch&LINK_LOC=blasthome). For prediction of the functional domain of FUST-1 NetNes (http://www.cbs.dtu.dk/services/NetNES/) and Prosite (http://prosite.expasy.org/scanprosite/) were used with FUST-1 coding sequence (Wormbase).

Paralysis assay

Worms were transferred on 5μM FUDR plates one day after L4. Worms were scored daily for movement for 12 days. Worms were counted as paralysed if they failed to move after

they were prodded on the nose. Experiments were performed at 20°C and at least 60 worms were counted per conditions. Survival curves and statistics were produced using Log-rank (Mantel-Cox) test. Standard error are shown on graph

Liquid culture assay

A synchronised population was obtained using hypochlorite extraction. Worms were grown on solid media up to day 1 of adulthood. At day 1, 30 worms per well were placed in S basal with OP-50 *E. coli* (optical density 0.5) in a flat-bottom 96-well plate. Measurement was done using Microtracker (Phylumtech), at least 3 wells were done per condition. Standard error are shown on graph

Neuronal integrity

To score gaps along the GABAergic neurons, day one, five and nine worms expressing *unc47p::mCherry* marker were selected for visualisation. To evaluate synaptic integrity, worms expressing *snb-1p::GFP* were selected. For neurodegeneration count during stress tests, adult day one worms were transferred to NGM + 400 mM NaCl at 20°C (osmotic stress) or normal NGM and put at 37°C (thermal stress) for six hours or for oxidative test worms were transferred on 240uM juglone for 30 minutes at 20°C. For visualization, animals were immobilized in M9 with 5 mM of levamisole and mounted on slides with 2% agarose pads. For all experiments, a minimum of 100 worms was scored for all

conditions. The mean and SEM were calculated and two-tailed t-tests were used for statistical analysis.

Lifespan assays

Worms were grown on NGM and transferred on NMG+ 5μM FUDR at day 1 of adulthood. Worms were counted every two days until their death. At least 100 worms were counted per strain. Survival curves and statistics were produced using Log-rank (Mantel-Cox) test. Standard error are shown on graph

Stress sensitivity assay

Worms were grown on NGM until day 1 of adulthood. At day 1, worms were transferred onto 400mM NaCl plates for osmotic stress, or 240 uM juglone for oxidative stress or onto NGM and put at 37°C for thermal stress. Worms were counted every two hours for up to 14 hours for osmotic and thermal stress and every 30 minutes for three hours for oxidative stress. Survival curves and statistics were produced using Log-rank (Mantel-Cox) test. Standard error are shown on graph

RT-PCR

Worms grown on NGM plates were collected before starvation and froze in Trizol. Total RNA was extracted according to the manufacturer protocol. 1ug of RNA was used to

produce cDNA using Vilo cDNA enzyme. Taqman probes detecting *fust-1* (probe Ce02434658_g1 Life technologies) and act-5 (probe Ce02454560_g1 Life technologies) as endogenous control were used. Expression level were calculated by converting the threshold cycle (Ct) values using the 2-ΔΔCt method (49)

ACKNOWLEDGEMENTS:

Some worm strains were provided by the CGC, which is funded by NIH Office of Research Infrastructure Programs (P40 OD010440), the International *C. elegans* Gene knock-out consortium and the TransgenOme project. We thank Sarah Peyrard for technical support.

Funding

M.T. is supported by Reseau de médecine généetique appliquée and by the Fonds de recherche du Québec- Santé , J.A.P. holds a career development award from the Fonds de recherche du Québec - Santé. This research was funded by the Canadian Institute of Health Research-CIHR and the ALS Society of Canada.

REFERENCES:

1. Vance,C., Rogelj,B., Hortobágyi,T., De Vos,K.J., Nishimura,A.L., Sreedharan,J., Hu,X., Smith,B., Ruddy,D., Wright,P., *et al.* (2009) Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science*, **323**, 1208–1211.
2. Kwiatkowski,T.J., Bosco,D.A., LeClerc,A.L., Tamrazian,E., Vanderburg,C.R., Russ,C., Davis,A., Gilchrist,J., Kasarskis,E.J., Munsat,T., *et al.* (2009) Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science*, **323**, 1205–1208.
3. Deng,H., Gao,K. and Jankovic,J. (2014) The role of FUS gene variants in neurodegenerative diseases. *Nat Rev Neurol*, 10.1038/nrneurol.2014.78.
4. Lagier-Tourenne,C., Polymenidou,M. and Cleveland,D.W. (2010) TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration. *Hum. Mol. Genet.*, **19**, R46–R64.
5. Hoell,J.I., Larsson,E., Runge,S., Nusbaum,J.D., Duggimpudi,S., Farazi,T.A., Hafner,M., Borkhardt,A., Sander,C. and Tuschl,T. (2011) RNA targets of wild-type and mutant FET family proteins. *Nature Structural & Molecular Biology*, **18**, 1428–1431.
6. Kovar,H. (2011) Dr. Jekyll and Mr. Hyde: The Two Faces of the FUS/EWS/TAF15 Protein Family. *Sarcoma*, **2011**, 837474–13.
7. Couthouis,J., Hart,M.P., Erion,R., King,O.D., Diaz,Z., Nakaya,T., Ibrahim,F., Kim,H.-J., Mojsilovic-Petrovic,J., Panossian,S., *et al.* (2012) Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. *Hum. Mol. Genet.*, **21**, 2899–2911.
8. Couthouis,J., Hart,M.P., Shorter,J., DeJesus-Hernandez,M., Erion,R., Oristano,R., Liu,A.X., Ramos,D., Jethava,N., Hosangadi,D., *et al.* (2011) A yeast functional screen predicts new candidate ALS disease genes. *Proceedings of the National Academy of Sciences*, **108**, 20881–20890.
9. Therrien,M. and Parker,J.A. (2014) Worming forward: amyotrophic lateral sclerosis toxicity mechanisms and genetic interactions in Caenorhabditis elegans. *Front Genet*, **5**, 1–13.
10. Xia,R., Liu,Y., Yang,L., Gal,J., Zhu,H. and Jia,J. (2012) Motor neuron apoptosis and neuromuscular junction perturbation are prominent features in a Drosophila model of Fus-mediated ALS. *Molecular Neurodegeneration*, **7**, 10.
11. Sasayama,H., Shimamura,M., Tokuda,T., Azuma,Y., Yoshida,T., Mizuno,T., Nakagawa,M., Fujikake,N., Nagai,Y. and Yamaguchi,M. (2012) Knockdown of the Drosophila fused in sarcoma (FUS) homologue causes deficient locomotive behavior and shortening of motoneuron terminal branches. *PLoS ONE*, **7**, e39483.
12. Shahidullah,M., Le Marchand,S.J., Fei,H., Zhang,J., Pandey,U.B., Dalva,M.B., Pasinelli,P. and Levitan,I.B. (2013) Defects in Synapse Structure and Function Precede Motor Neuron Degeneration in Drosophila Models of FUS-Related ALS. *Journal of Neuroscience*, **33**, 19590–19598.
13. Frickenhaus,M., Wagner,M., Mallik,M., Catinozzi,M. and Storkebaum,E. (2015) Highly efficient cell-type-specific gene inactivation reveals a key function for the Drosophila FUS homolog cabeza in neurons. *Sci. Rep.*, **5**, 9107.
14. Therrien,M., Rouleau,G.A., Dion,P.A. and Parker,J.A. (2013) Deletion of C9ORF72

results in motor neuron degeneration and stress sensitivity in C. elegans. *PLoS ONE*, **8**, e83450.
15. Vaccaro,A., Tauffenberger,A., Aggad,D., Rouleau,G., Drapeau,P. and Parker,J.A. (2012) Mutant TDP-43 and FUS cause age-dependent paralysis and neurodegeneration in C. elegans. *PLoS ONE*, **7**, e31321.
16. Maruyama,I.N. and Brenner,S. (1991) A phorbol ester/diacylglycerol-binding protein encoded by the unc-13 gene of Caenorhabditis elegans. *Proc. Natl. Acad. Sci. U.S.A.*, **88**, 5729–5733.
17. Therrien,M., Rouleau,G.A., Dion,P.A. and Parker,J.A. (2013) Deletion of C9ORF72 Results in Motor Neuron Degeneration and Stress Sensitivity in C. elegans. *PLoS ONE*, **8**, e83450.
18. Vaccaro,A., Patten,S.A., Ciura,S., Maios,C., Therrien,M., Drapeau,P., Kabashi,E. and Parker,J.A. (2012) Methylene blue protects against TDP-43 and FUS neuronal toxicity in C. elegans and D. rerio. *PLoS ONE*, **7**, e42117.
19. Simonetta,S.H. and Golombek,D.A. (2007) An automated tracking system for Caenorhabditis elegans locomotor behavior and circadian studies application. *J. Neurosci. Methods*, **161**, 273–280.
20. Barbagallo,B., Prescott,H.A., Boyle,P., Climer,J. and Francis,M.M. (2010) A dominant mutation in a neuronal acetylcholine receptor subunit leads to motor neuron degeneration in Caenorhabditis elegans. *J. Neurosci.*, **30**, 13932–13942.
21. Jorgensen,E.M. (2005) GABA. *WormBook*, 10.1895/wormbook.1.14.1.
22. Jin,Y. (2005) Synaptogenesis. *WormBook*, 10.1895/wormbook.1.44.1.
23. Mahoney,T.R., Luo,S. and Nonet,M.L. (2006) Analysis of synaptic transmission in Caenorhabditis elegans using an aldicarb-sensitivity assay. *Nat Protoc*, **1**, 1772–1777.
24. Vashlishan,A.B., Madison,J.M., Dybbs,M., Bai,J., Sieburth,D., Ch'ng,Q., Tavazoie,M. and Kaplan,J.M. (2008) An RNAi screen identifies genes that regulate GABA synapses. *Neuron*, **58**, 346–361.
25. Saifee,O., Wei,L. and Nonet,M.L. (1998) The Caenorhabditis elegans unc-64 locus encodes a syntaxin that interacts genetically with synaptobrevin. *Mol. Biol. Cell*, **9**, 1235–1252.
26. Murphy,C.T. and Hu,P.J. (2013) Insulin/insulin-like growth factor signaling in C. elegans. *WormBook*, 10.1895/wormbook.1.164.1.
27. de Castro,E., Hegi de Castro,S. and Johnson,T.E. (2004) Isolation of long-lived mutants in Caenorhabditis elegans using selection for resistance to juglone. *Free Radical Biology and Medicine*, **37**, 139–145.
28. Murphy,C.T., McCarroll,S.A., Bargmann,C.I., Fraser,A., Kamath,R.S., Ahringer,J., Li,H. and Kenyon,C. (2003) Genes that act downstream of DAF-16 to influence the lifespan of Caenorhabditis elegans. *Nature*, **424**, 277–283.
29. Oliveira,R.P., Porter Abate,J., Dilks,K., Landis,J., Ashraf,J., Murphy,C.T. and Blackwell,T.K. (2009) Condition-adapted stress and longevity gene regulation by Caenorhabditis elegans SKN-1/Nrf. *Aging Cell*, **8**, 524–541.
30. Broughton,S. and Partridge,L. (2009) Insulin/IGF-like signalling, the central nervous system and aging. *Biochem. J.*, **418**, 1–12.
31. Boccitto,M., Lamitina,T. and Kalb,R.G. (2012) Daf-2 signaling modifies mutant SOD1 toxicity in C. elegans. *PLoS ONE*, **7**, e33494.

32. Teixeira-Castro,A., Ailion,M., Jalles,A., Brignull,H.R., Vilaça,J.L., Dias,N., Rodrigues,P., Oliveira,J.F., Neves-Carvalho,A., Morimoto,R.I., *et al.* (2011) Neuron-specific proteotoxicity of mutant ataxin-3 in C. elegans: rescue by the DAF-16 and HSF-1 pathways. *Hum. Mol. Genet.*, **20**, 2996–3009.
33. Vaccaro,A., Tauffenberger,A., Ash,P.E.A., Carlomagno,Y., Petrucelli,L. and Parker,J.A. (2012) TDP-1/TDP-43 regulates stress signaling and age-dependent proteotoxicity in Caenorhabditis elegans. *PLoS Genet*, **8**, e1002806.
34. Zhang,T., Baldie,G., Periz,G. and Wang,J. (2014) RNA-Processing Protein TDP-43 Regulates FOXO-Dependent Protein Quality Control in Stress Response. *PLoS Genet*, **10**, e1004693.
35. Zhang,T., Hwang,H.Y., Hao,H., Talbot,C. and Wang,J. (2012) Caenorhabditis elegans RNA-processing Protein TDP-1 Regulates Protein Homeostasis and Life Span. *Journal of Biological Chemistry*, **287**, 8371–8382.
36. Liu-Yesucevitz,L., Bilgutay,A., Zhang,Y.-J., Vanderweyde,T., Vanderwyde,T., Citro,A., Mehta,T., Zaarur,N., McKee,A., Bowser,R., *et al.* (2010) Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue. *PLoS ONE*, **5**, e13250.
37. Colombrita,C., Zennaro,E., Fallini,C., Weber,M., Sommacal,A., Buratti,E., Silani,V. and Ratti,A. (2009) TDP-43 is recruited to stress granules in conditions of oxidative insult. *Journal of Neurochemistry*, **111**, 1051–1061.
38. Gal,J., Zhang,J., Kwinter,D.M., Zhai,J., Jia,H., Jia,J. and Zhu,H. (2011) Nuclear localization sequence of FUS and induction of stress granules by ALS mutants. *Neurobiology of Aging*, **32**, 2323.e27–40.
39. Bosco,D.A., Lemay,N., Ko,H.K., Zhou,H., Burke,C., Kwiatkowski,T.J., Sapp,P., McKenna-Yasek,D., Brown,R.H. and Hayward,L.J. (2010) Mutant FUS proteins that cause amyotrophic lateral sclerosis incorporate into stress granules. *Hum. Mol. Genet.*, **19**, 4160–4175.
40. Sama,R.R.K., Ward,C.L., Kaushansky,L.J., Lemay,N., Ishigaki,S., Urano,F. and Bosco,D.A. (2013) FUS/TLS assembles into stress granules and is a prosurvival factor during hyperosmolar stress. *J. Cell. Physiol.*, 10.1002/jcp.24395.
41. Andersson,M.K., Ståhlberg,A., Arvidsson,Y., Olofsson,A., Semb,H., Stenman,G., Nilsson,O. and Aman,P. (2008) The multifunctional FUS, EWS and TAF15 proto-oncoproteins show cell type-specific expression patterns and involvement in cell spreading and stress response. *BMC Cell Biol.*, **9**, 37.
42. Blechingberg,J., Luo,Y., Bolund,L., Damgaard,C.K. and Nielsen,A.L. (2012) Gene expression responses to FUS, EWS, and TAF15 reduction and stress granule sequestration analyses identifies FET-protein non-redundant functions. *PLoS ONE*, **7**, e46251.
43. Wang,J.-W., Brent,J.R., Tomlinson,A., Shneider,N.A. and McCabe,B.D. (2011) The ALS-associated proteins FUS and TDP-43 function together to affect Drosophila locomotion and life span. *J. Clin. Invest.*, **121**, 4118–4126.
44. Kino,Y., Washizu,C., Kurosawa,M., Yamada,M., Miyazaki,H., Akagi,T., Hashikawa,T., Doi,H., Takumi,T., Hicks,G.G., *et al.* (2015) FUS/TLS deficiency causes behavioral and pathological abnormalities distinct from amyotrophic lateral sclerosis. *Acta Neuropathol Commun*, **3**, 24.

45. Zhou,Y., Liu,S., Liu,G., Oztürk,A. and Hicks,G.G. (2013) ALS-associated FUS mutations result in compromised FUS alternative splicing and autoregulation. *PLoS Genet*, **9**, e1003895.
46. Merner,N.D., Girard,S.L., Catoire,H., Bourassa,C.V., Belzil,V.V., Rivière,J.-B., Hince,P., Levert,A., Dionne-Laporte,A., Spiegelman,D., *et al.* (2012) Exome sequencing identifies FUS mutations as a cause of essential tremor. *Am. J. Hum. Genet.*, **91**, 313–319.
47. Stiernagle,T. (2006) Maintenance of C. elegans. *WormBook*, 10.1895/wormbook.1.101.1.
48. Sarov,M., Schneider,S., Pozniakovski,A., Roguev,A., Ernst,S., Zhang,Y., Hyman,A.A. and Stewart,A.F. (2006) A recombineering pipeline for functional genomics applied to Caenorhabditis elegans. *Nat Meth*, **3**, 839–844.
49. Livak,K.J. and Schmittgen,T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*, **25**, 402–408.

FIGURE

Figure 1

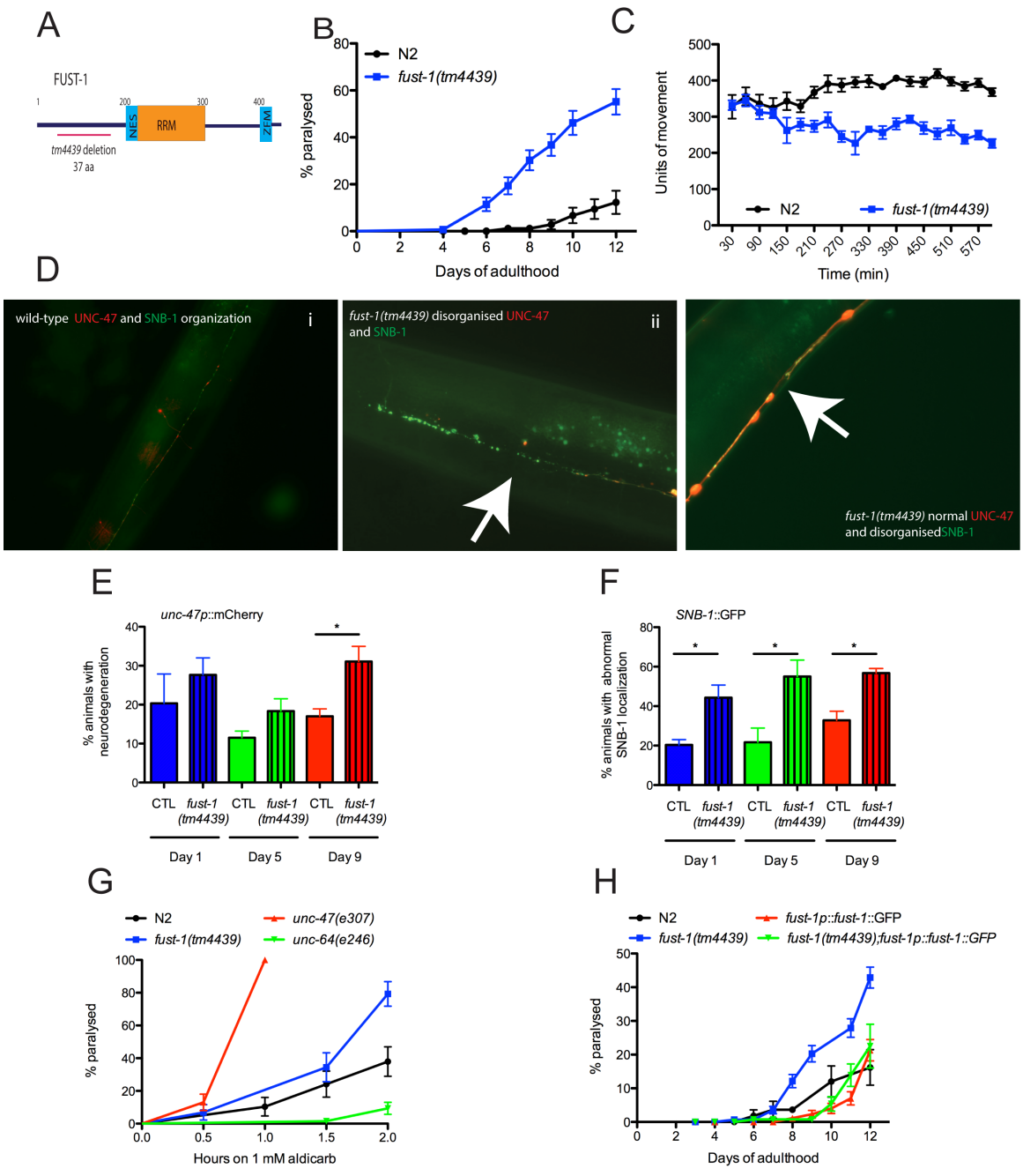


Figure AI.1. Deletion of *fust-1* causes motility impairment and loss of neuronal integrity. (A) *C. elegans fust-1* is the ortholog of human *FUS*, *EWSR1* and *TAF15* and contains the conserved RNA recognition motif (RRM), nuclear export signal (NES) and zinc finger motif (ZFM) **(b-c)** Loss of *fust-1* expression causes age-dependant paralysis **(B)** on solid media (p value <0.0001) and **(C)** the phenomenon is accelerated in liquid media (p value <0.0001) **(D-F) (D)** *unc-47* promoter (red) and SNB-1 (green) were used to visualise motor neurons and synapse formation(normal expression i) Decreased expression of *fust-1* caused gaps along the motor neurons (ii) and disorganisation of the SNB-1 protein (ii-iii). Quantification of the number of animals at day 1, 5 and 9 with **(E)** neurodegeneration in motor neuron (*unc-47p*::mCherry) (*p value <0.05, n≥100 for each conditions) and **(F)** synaptic disorganisation (*unc25p*::SNB-1::GFP) (*p value <0.05, n≥100 for each conditions). **(G)** *fust-1(tm4439)* mutants are more sensitive to aldicarb than N2 and *unc-64(e246)* but less than the hypersensitive strain *unc-47(307)* p<0.001 . **(H)** Overexpression of *fust-1* rescues the age dependant paralysis phenotype observed in *fust-1(tm4439)* mutants (p value <0.001)

Figure 2

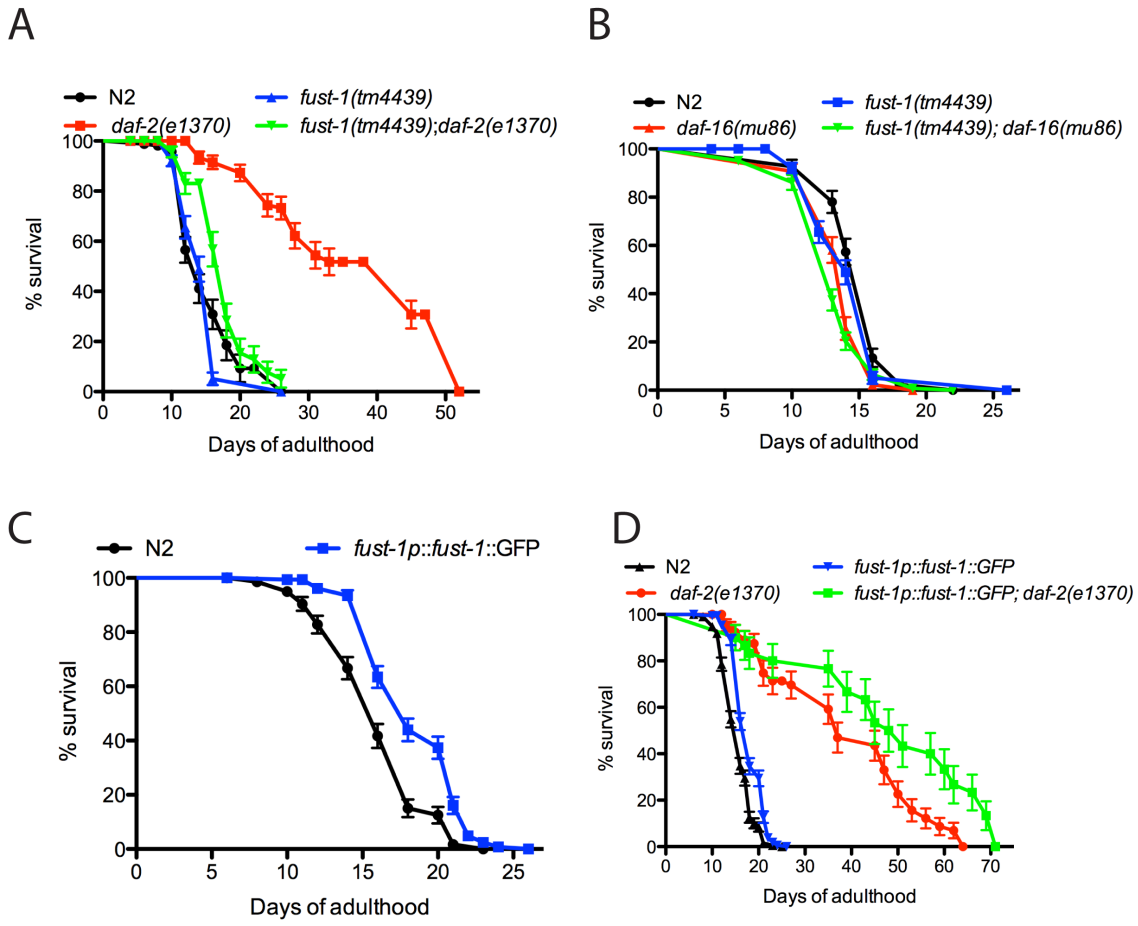


Figure AI.2. *fust-1* functions within the IIS pathway to regulate lifespan. (A) Decreased expression of *fust-1* abolished the long-lived phenotype of *daf-2(e1370)* (p value <0.005) but **(B)** has no effect on *daf-16* mutants. Overexpression of *fust-1* increased the lifespan **(C)** of N2 worms (p value <0.0001) and **(D)** of *daf-2(e1370)* mutants (p value <0.005)

Figure 3

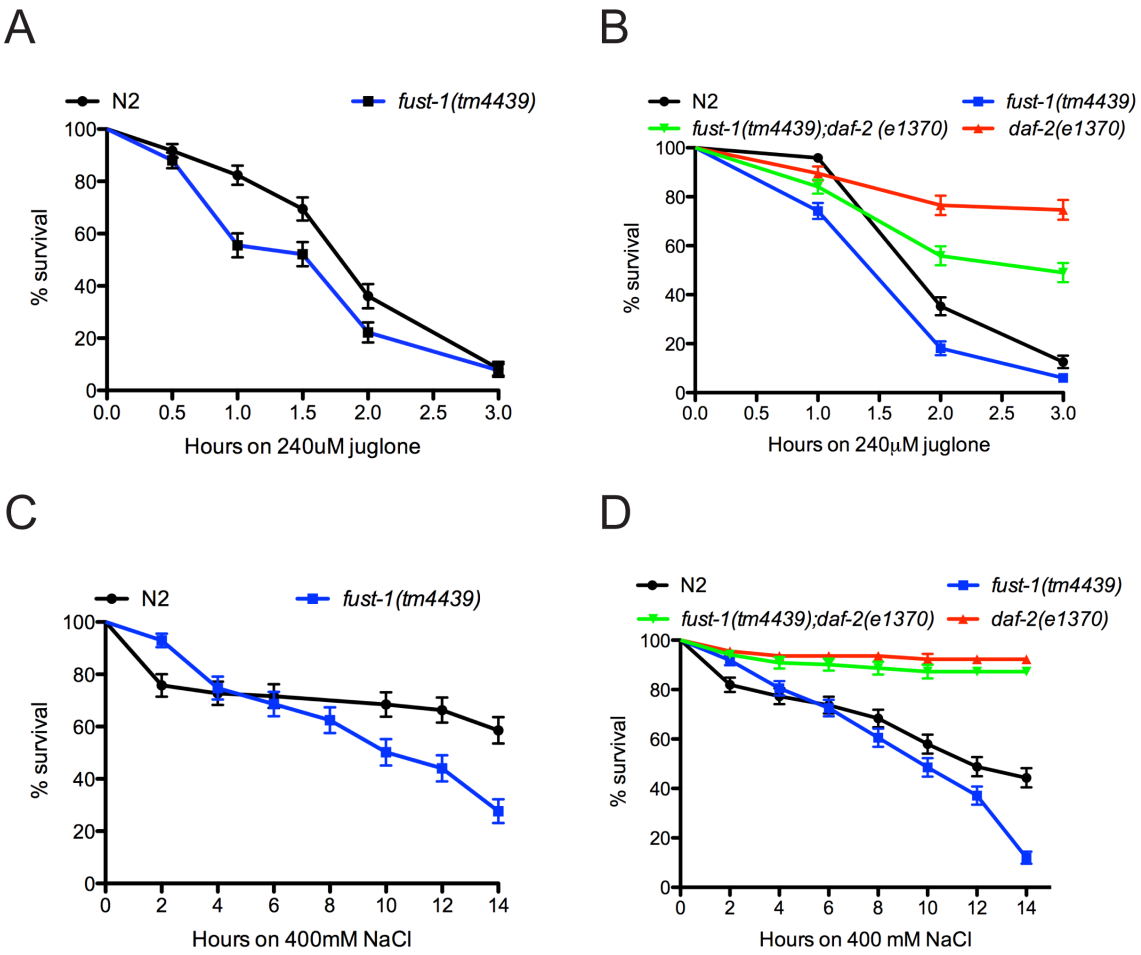


Figure AI.3. Oxidative and osmotic stress responses require *fust-1* (A-B) (A) *fust-1* mutants are more sensitive to oxidative stress induced by juglone compared to N2 controls (p value <0.0005). **(B)** *fust-1; daf-2* mutants are more sensitive to oxidative stress than *daf-2* controls (p value <0.0001. **(C-D) (C)** *fust-1* mutants are sensitive to osmotic stress (p value <0.001) but the *fust-1* mutation does affect the stress sensitivity of **(D)** *daf-2(e1370)*.

Figure 4

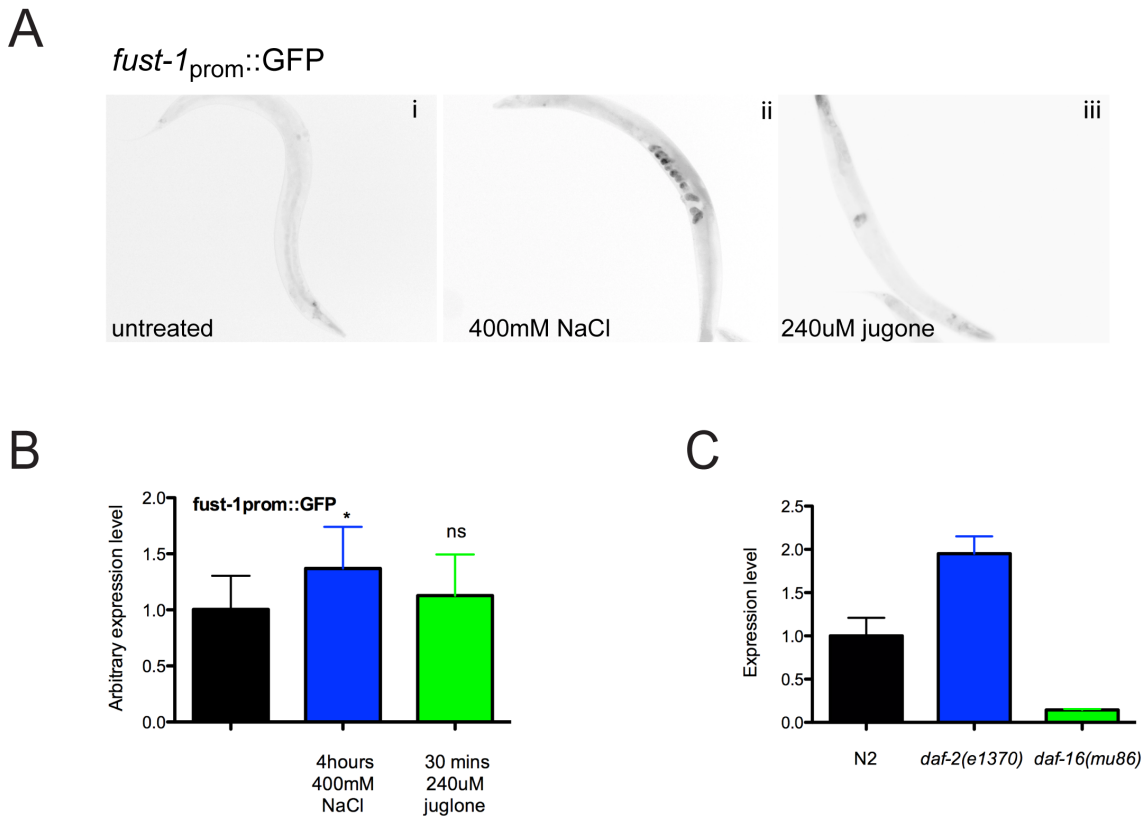


Figure AI.4. *fust-1* expression is induced by osmotic stress and IIS. (A) Representative, black and white, photo-reversed images of the transgenic *fust-1p*::GFP reporter strains (control i) showing increased expression in response to osmotic (ii), but not oxidative stress (iii). **(B)** Relative quantification level of *fust-1p*::GFP under stress conditions (* p value <0.0001, n≥30 for each condition). **(C)** qRT-PCR with ΔΔCT analysis of *fust-1* expression showing an increased expression in *daf-2(e1370)* mutants and decreased expression in *daf-16(mu86)* mutants.

Figure 5

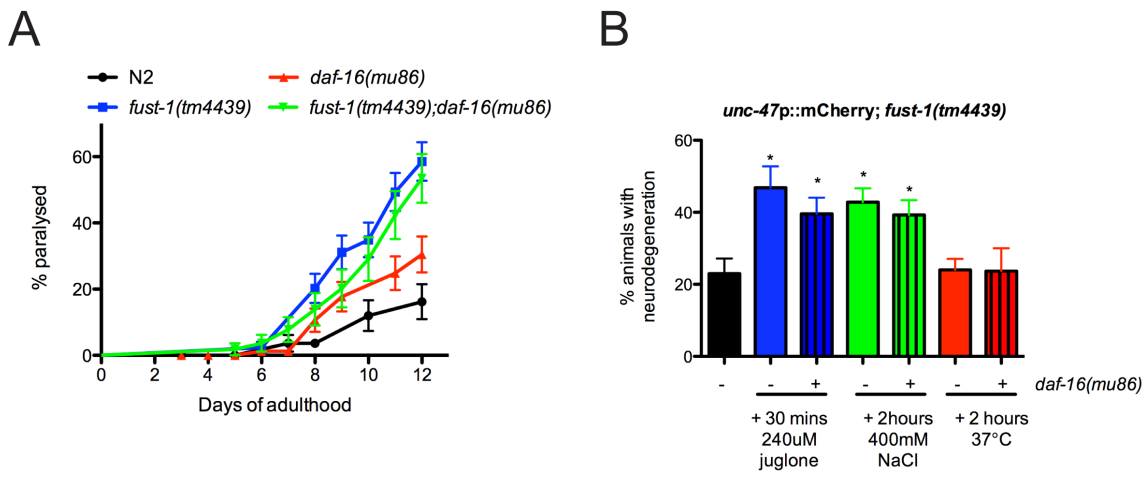
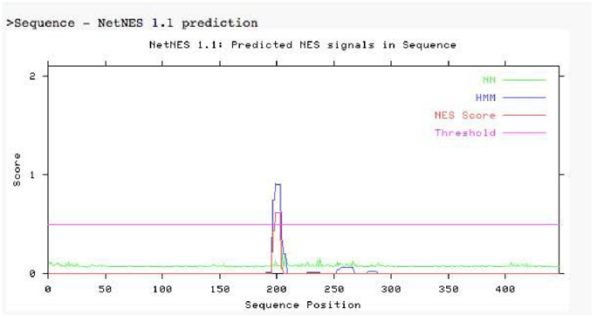


Figure AI.5. Maintenance of neuronal integrity by *fust-1* is not regulated by the IIS. **(A)** *daf-16(mu86)*; *fust-1(tm443)* mutants had rates of paralysis similar to *fust-1(tm443)* mutants alone. **(B)** Acute osmotic and oxidative stresses induce neurodegeneration in day 1 *fust-1* mutants but not thermal stress. *daf-16* mutants do not influence the % of animals with neurodegeneration (*p value < 0.05, n≥60 for each conditions).

SUPPLEMENTARY MATERIAL

A



B

C

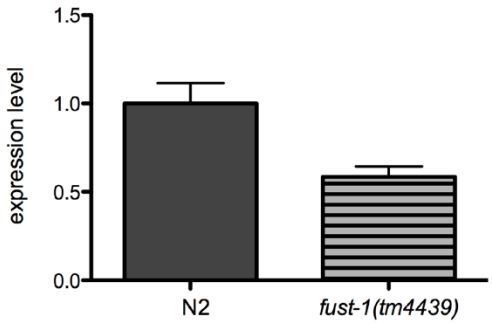


Figure AI.S1. FUST-1 has the functional domains of human FUS, EWSR1 and TAF15. **(A)** NetNES prediction showing the presence of a potential NES at amino acid 200 of FUST-1. **(B)** Prosite prediction of FUST-1 showing the presence of an RNA-recognition motif and a zinc-finger motif. **(C)** qRT-PCR with ΔΔCT analysis of *fust-1* expression showing that *fust-1(tm4439)* exhibit 50% expression when compared to wild-type N2 worms.

Figure AI.S2. FUST-1 overexpression strain (A) Picture of *fust-1p*::*fust-1*::GFP strain showing expression of *fust-1* in head, pharynx, intestine and tail of the adult animal. Left panel is showing GFP image merged with DIC image and right image is black and white inversion of the same image.

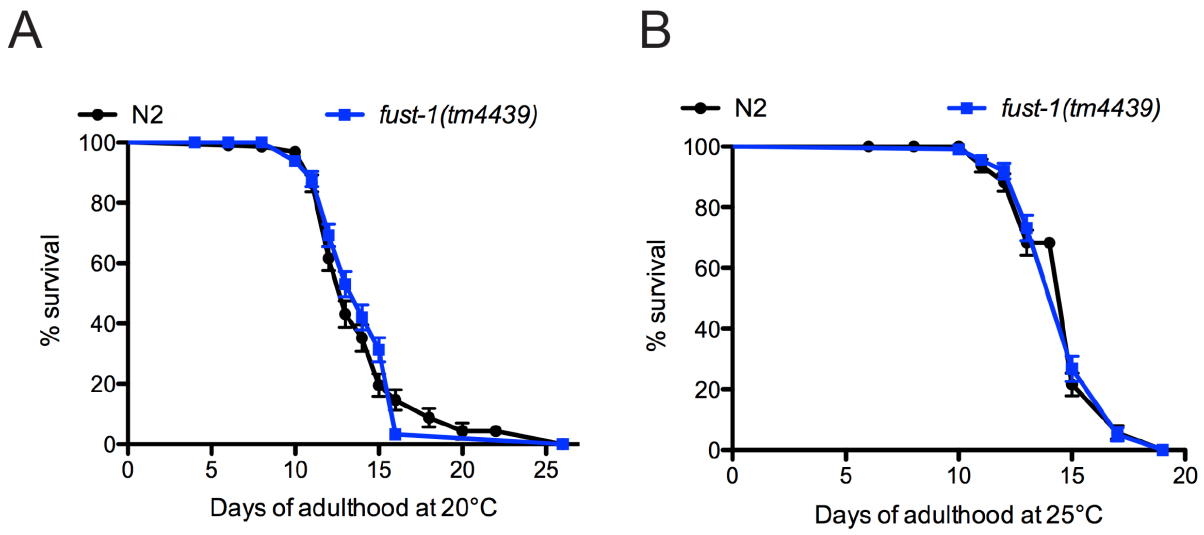


Figure AI.S3. Loss of *fust-1* does not affect lifespan. *fust-1(tm4439)* mutants had lifespans similar to N2 controls when grown at either **(A)** 20°C or **(B)** 25°C.

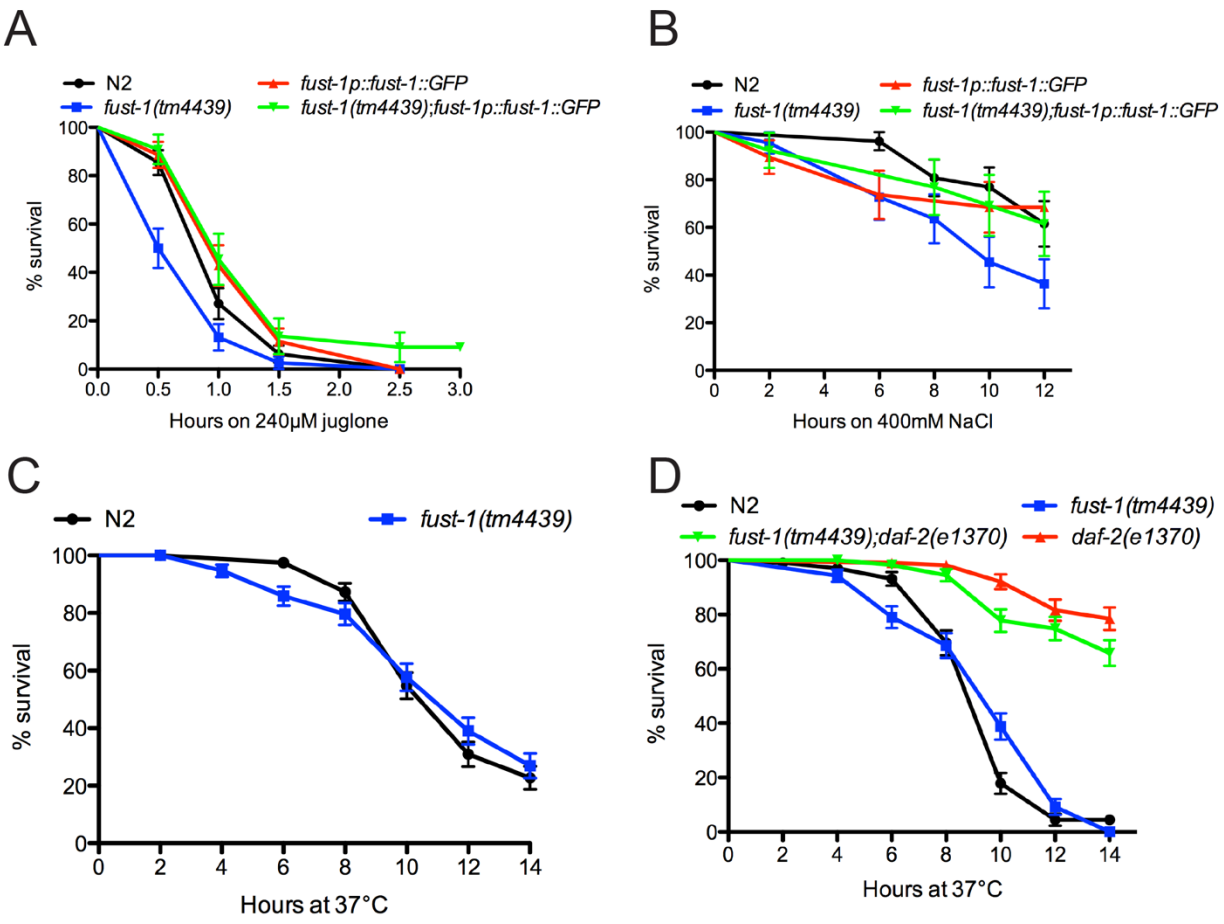


Figure AI.S4. Overexpression of *fust-1* rescues stress sensitivity of *fust-1(tm4439)*. Overexpression of *fust-1* **(A)** rescues *fust-1(tm4439)* sensitivity to juglone (p value <0.001) and **(B)** partially rescues sensitivity to osmotic stress (p value< 0.16). **(C-D)** Decreased expression of *fust-1* does not affect response to thermal stress of wild-type and *daf-2* mutants.

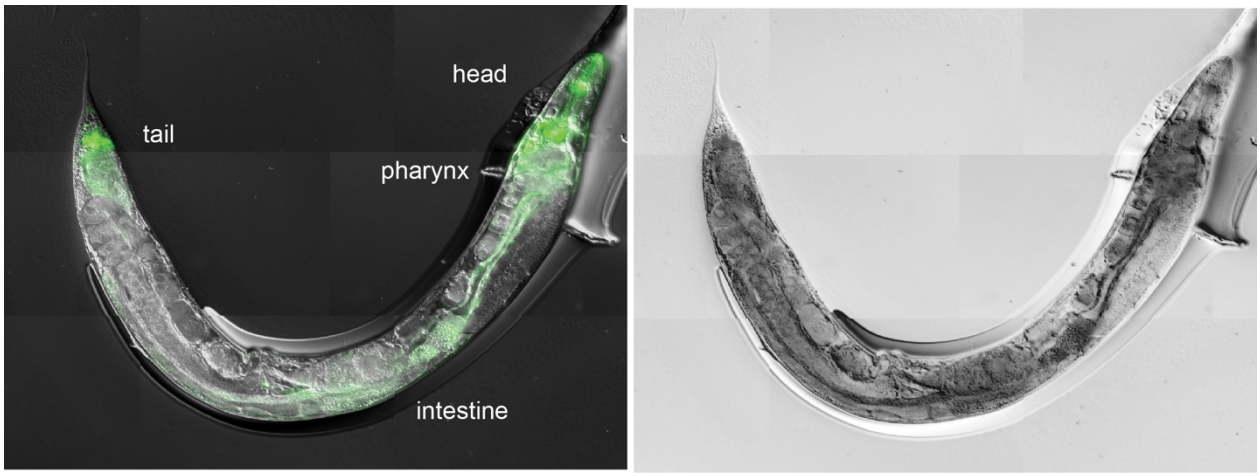


Figure AI.S5. Representative picture of *fust-1p*::GFP worm. Picture of *fust-1p*::GFP strain showing expression of *fust-1* in head, pharynx, intestine and tail of the adult animal. Left panel is showing GFP and DIC image and right panel is the black and white inversion of the same image.

Table AI.S1: Statistic tables of figure AI.1

A) paralysis on solid media of *fust-1(tm4439)*

	% paralysed at day 12	P value vs N2
N2	13%	---
fust-1(tm4439)	66%	<0.0001

B)paralysis in liquid culture of *fust-1(tm4439)*

	P value vs N2
N2	---
fust-1(tm4439)	<0.0001

C) *fust-1prom*::FUST-1::GFP rescue *fust-1(tm4439*) motor phenotype

	% paralysed at day 12	P value vs N2	P value vs *fust-1 (tm4439)*
N2	16%	---	0.0010
fust-1(tm4439)	43%	0.0010	-----
fust-1p::fust-1::GFP	21%	0.5907	<0.0001
fust-1(tm4439); fust-1p::fust-1::GFP	22%	0.6629	0.0006

Table AI.S2: Statistic tables of figure AI.2

A) *fust-1(tm4439)* decreased long-lived phenotype of *daf-2(e1370)* mutants

	Median survival	P value vs N2	P value vs *fust-1 (tm4439)*
N2	14 days	---	0.4206
fust-1(tm4439)	14 days	0.4206	---
daf-2(e1370)	45 days	<0.0001	<0.0001
fust-1(tm4439);daf-2(e1370)	18 days	0.003	<0.0001

B) *fust-1(tm4439)* does not affect *daf-16* mutants lifespan

	Median survival	P value vs N2	P value vs *fust-1 (tm4439)*
N2	16 days	---	0.0761
fust-1(tm4439)	14 days	0.0761	-----
daf-16(mu86)	14 days	<0.0001	0.0500
fust-1(tm4439);daf 16(mu86)	13 days	<0.0001	0.0072

C) *fust-1(tm4439)* has no effect on lifespan at 20°C

At 20°C	Median survival	P value vs N2
N2	13 days	---
fust-1(tm4439)	14 days	0.5728

D) *fust-1(tm4439)* has no effect on lifespan at 25°C

At 25°C	Median survival	P value vs N2
N2	15 days	---
fust-1(tm4439)	15 days	0.3910

E) *fust*-1 overexpression increases lifespan

At 20°C	Median survival	P value vs N2
N2	16 days	---
*fust-1p::fust-1::*GFP	18 days	<0.0001

F) *fust-1* overexpression increases *daf-2(e1370)* lifespan

	Median survival	P value vs N2	P value vs *fust-1prom*::FUST-1::GFP
N2	16 days	---	<0.0001
fust-1p::fust-1::GFP	18 days	<0.0001	-----
daf2(e1370)	37 days	<0.0001	<0.0001
fust-1p::fust-1::GFP;daf-2(e1370)	49.5 days	<0.0001	<0.0001

Table AI.S3 –Statistic tables of figure AI.5

	% paralysed at day 12	P value vs N2	P value vs *fust-1 (tm4439)*
N2	16%	---	<0.0001
fust-1(tm4439)	58%	<0.0001	-----
daf-16(mu86)	30%	0.0708	0.0008
fust-1(tm4439); daf-16(mu86)	53%	<0.0001	0.5059

Table AI.S4 - List of strains

Strain	Source	Genotype
N2	CGC	Wild-type
FX4439	National Bioresource Project of Japan	*fust-1(tm4439)*
ufls34	M. Francis lab	*ufls34* (unc-47::mCherry)
CZ333	CGC	unc25p::snb-1::GFP
CB307	CGC	*unc-47(e307)*
CB246	CGC	*unc-64(e246)*
CB1370	CGC	*daf-2(e1370)*
CF1038	CGC	*daf-16(mu86)*
XQ 307	JA. Parker lab	*unc-119(ed3);sEx307* (fust-1p::fust-1::GFP)
BC10929	CGC	*dpy-5(e907)*I;*sEx10929*(rCes C27H5.3::GFP + pCeh361)