Genetics of Amyotrophic Lateral Sclerosis: A Review

Intan Barul Akma Bakhtiar a,b* a Neurogenetics Group, Department of Brain Sciences, Faculty of Medicine, Imperial College London, London, UK.
b Faculty of Allied Health Sciences, University of Cyberjaya, Persiaran Bestari 1, Cyber 11, 63000 Cyberjaya, Selangor, Malaysia.
*Corresponding author.
Email address of corresponding author: intan.akma@cyberjaya.edu.my

ABSTRACT

Amyotrophic lateral sclerosis (ALS) is a fatal neurodegenerative disorder with a poor prognosis that affects both upper motor neurons (UMN) and lower motor neurons (LMN). Currently, 126 candidate genes associated with familial and sporadic ALS have been identified. The exact pathogenesis mechanisms associated with ALS remains unknown however, several underlying ALS pathogenesis resulting in motor neuron degeneration have been identified including superoxide metabolism, RNA metabolism, autophagy, endoplasmic reticulum (ER) stress and the unfolded protein response (UPR), cell division, axonal transport, D-amino acid degradation and apoptosis. Here, we present a comprehensive review the genetics of ALS, in particular, the most commonly mutated ALS-related genes which may provide insightful information to understand their involvement in ALS.

INTRODUCTION

Amyotrophic lateral sclerosis (ALS), also known as Lou Gehrig’s disease, is a fatal neurodegenerative disease without effective treatments. It is also the most common adult-onset motor neuron disorder. The disease is characterised by the loss of upper motor neurons in motor cortex and lower motor neurons in the brainstem and spinal cord, that results in progressive paralysis and death due to respiratory failure, within three years of onset of the first symptoms (1). ALS affects people worldwide and although the exact incidence is unknown, ALS incidence across the European populations is fairly uniform, with a median incidence rate is 2.08 per 100 000 persons per year and a median prevalence rate of 5.40 per 100 000 persons per year (2). In sporadic ALS (SALS), the incidence is higher in men (3.0 per 100 000 person-years) than in women (2.4 per 100 000 person-years), however, in familial ALS (FALS) the incidence in men and women is similar (3). The disease usually occurs after 40 years of age, between 47 to 63 years of age and incidence of ALS diminishes rapidly after the age of 80 (3). Limb-onset ALS, is identified by the coexistence of upper motor neuron (UMN) and lower motor
neuron (LMN) signs in the limbs, is the most predominant form of ALS (4), occurring in 70% of cases, whereas 25% of cases display bulbar-onset ALS which is characterised by spastic dysarthria (slurring of speech) and dysphagia (difficulty in swallowing) with subsequent development of limb features (4).

ALS is categorised into familial and sporadic ALS, which are clinically indistinguishable. SALS accounts for the majority of ALS cases, whereas about 10% of ALS cases are familial (4). However it has been reported that gene mutations occurring in FALS cases are also found to be the causative gene mutations in SALS (5). It has also been observed that in the Irish population, the first-degree relatives of SALS patients may have a higher risk of ALS and other neurodegenerative diseases compared to controls (6). ALS is also associated with frontotemporal degeneration (FTD) in which up to 50% of ALS patients display cognitive deficits (7, 8). There is no significant difference in the level of cognitive impairment between SALS and FALS patients (9), whereas 14% of FTD patients are diagnosed with ALS (10). Signs of frontotemporal lobe deterioration include language, executive dysfunction and behavioural changes that include irritability, confusion and diminished empathy (11). Neuropathologically, TAR DNA binding protein-43 (TBP-43) and fused in sarcoma (FUS) protein are also present in ubiquitin-positive cytoplasmic inclusions of ALS and FTD patients, supporting the common pathogenesis of these two diseases (12).

ALS has a poor prognosis with about 20% of ALS patients surviving between 5 to 10 years after symptom onset with the respiratory phenotype exhibited the worst prognosis (13). Prolonged survival indicators are limb-onset disease and younger age at onset whereas bulbar-onset disease, early respiratory muscle dysfunction and older age at onset indicate reduced survival (14). Currently, Riluzole is the only widely available disease-modifying drug for the treatment of ALS however, it only prolongs survival by two to three months (15). Riluzole blocks voltage-gated sodium channel therefore reducing the release of glutamate (16). Edaravone (Radicava), a free radical scavenger has been approved in Japan and the US for ALS treatment. In clinical trials in Japan, ALS patients receiving Edaravone showed reduced loss of physical function as compared to those receiving placebo (17, 18). Edaravone works by suppressing the formation of free radicals (19).
Fig. 1: Clinical manifestations of ALS. ALS is a neurodegenerative disease that is characterized by the loss of both upper motor neurons (UMN) and lower motor neurons (LMN). Muscle weakness and dysphagia are the main presentations of ALS but up to 50% of ALS patients have non-motor symptoms for example, cognitive impairment.

GENETICS OF ALS

The majority of familial ALS (FALS) cases show a Mendelian pattern of inheritance. Sporadic ALS cases are defined as having no relatives that are affected by ALS (20). Familial ALS (FALS) accounts for approximately 5 to 10% of ALS cases whilst the rest of ALS cases are sporadic (20) and in more than 80% of ALS cases, the responsible genes for ALS still remain largely unknown (5, 21, 22). Therefore, there is an important need to elucidate additional ALS causative genes which have been identified for ALS (http://alsod.iop.kcl.ac.uk). To date, 126 causative genes have been identified as being the most common in C9ORF72 (4, 23, 24). Other known genes include SOD1, TDP-43, FUS, and unknown genes.
mutations that cause a typical ALS phenotype are found in SOD1, TARDBP (TDP-43), FUS (TLS), ANG and OPTN (20). A few selected causative genes that are associated with ALS are summarised in Table 2 overleaf.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Locus</th>
<th>Inheritance</th>
<th>Putative protein function(s)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>SOD1</td>
<td>21q22</td>
<td>AD and AR</td>
<td>Superoxide metabolism</td>
<td>(25)</td>
</tr>
<tr>
<td>TARDBP</td>
<td>1p36</td>
<td>AD</td>
<td>RNA metabolism</td>
<td>(26)</td>
</tr>
<tr>
<td>FUS</td>
<td>16p11</td>
<td>AD and AR</td>
<td>RNA metabolism</td>
<td>(27)</td>
</tr>
<tr>
<td>C9ORF72</td>
<td>9p21</td>
<td>AD</td>
<td>DENN protein</td>
<td>(23)</td>
</tr>
<tr>
<td>VCP</td>
<td>9p13</td>
<td>AD</td>
<td>Proteasome and vesicle trafficking</td>
<td>(28)</td>
</tr>
<tr>
<td>SQSTM1</td>
<td>5q35</td>
<td>AD</td>
<td>Ubiquitination and autophagy</td>
<td>(29)</td>
</tr>
<tr>
<td>UBQLN2</td>
<td>Xp11</td>
<td>XD</td>
<td>Autophagy adaptor</td>
<td>(30)</td>
</tr>
<tr>
<td>OPTN</td>
<td>10p13</td>
<td>AR and AD</td>
<td>Autophagy adaptor</td>
<td>(31)</td>
</tr>
<tr>
<td>ANG</td>
<td>14q11</td>
<td>AD</td>
<td>Ribonuclease</td>
<td>(32)</td>
</tr>
<tr>
<td>VAPB</td>
<td>20q13.33</td>
<td>AD</td>
<td>Type IV membrane protein</td>
<td>(33)</td>
</tr>
<tr>
<td>ALS2</td>
<td>2q33-35</td>
<td>AR</td>
<td>Alsin Rho guanine nucleotide exchange factor</td>
<td>(34)</td>
</tr>
<tr>
<td>SETX</td>
<td>9q34</td>
<td>AD</td>
<td>RNA metabolism and probable RNA/DNA helicase</td>
<td>(35)</td>
</tr>
<tr>
<td>DCTN1</td>
<td>2p13</td>
<td>AD</td>
<td>Cell division and axonal transport</td>
<td>(36)</td>
</tr>
<tr>
<td>DAO</td>
<td>12q22-23</td>
<td>AD</td>
<td>Enzyme D-amino acid oxidase for D-amino acids degradation</td>
<td>(37)</td>
</tr>
</tbody>
</table>

Table 2: Some of causative genes for ALS. (Abbreviations: AD, autosomal dominant; AR, autosomal recessive; XD, X-linked dominant; DENN, differentially expressed in normal and neoplasia)
Superoxide dismutase 1 (SOD1)

Superoxide dismutase 1 (SOD1) encodes for a homodimeric metalloenzyme that catalyses the reaction that changes toxic $O_2^-$ into $O_2$ and $H_2O_2$, which is also known as the ‘Fenton reaction’. Approximately 12% of FALS and 1% of SALS cases are attributed to SOD1 mutations (22). Currently, 186 mutations in the SOD1 gene were reported in the ALSoD database (http://alsod.iop.kcl.ac.uk). SOD1 mutations are associated with the earlier age of onset than in SALS and longer duration of disease, usually between 5 to 10 years (38, 39). The A4V mutation in SOD1 which is the most frequent variant in the United States (40) is also the most aggressive form of SOD1 mutations with an average survival rate of two years after initial symptom onset and pure lower motor neuron (LMN) symptoms (39). The A4V mutation however, is rare in the UK. Misfolded superoxide dismutase 1 (SOD1) protein is colocalized in cytoplasmic neuronal inclusions in FALS and SALS motor neurons (41, 42).

TAR DNA-binding protein (TARDBP)

TAR DNA-binding protein (TARDBP) mutations are observed in approximately ~4% cases of FALS and a much smaller number of familial FTD cases and less than 1% of SALS cases worldwide (22, 43). Ribonucleoprotein binding and splicing are usually affected in these TARDBP mutations (22). TARDBP encodes for the transactive response (TAR) DNA binding protein 43 (TDP-43) that is present in the ubiquitinated neuronal cytoplasmic inclusions (NCIs) of all SALS patients and patients with frontotemporal degeneration with TDP-43 pathology (FTD-TDP), which is the most common form of FTD (44, 45). However, co-localisation of TDP-43 inclusions with SOD1 aggregates are rare in SOD1 SALS and are not present in FALS cases with SOD1 mutations (46). Longer disease duration (>60 months) is associated with TARDBP mutations in both Caucasian and Asian populations, however upper limb onset is more predominant in the Caucasian population whereas bulbar onset is more predominant in the Asian population (38). TDP-43 is also important in regulating endosomal trafficking and loss of TDP-43 function affects dendritic endosomes leading to decreased neuronal health (47).

Fused in sarcoma (FUS)

Another gene that is involved in RNA processing and splicing is fused in sarcoma (FUS), which encodes for translated in liposarcoma (TLS) protein. Like TDP-43, FUS protein binds RNA at the C-terminus (22). FUS directly binds to single-and double-stranded DNA (48, 49) and is implicated in transcriptional regulation and DNA damage repair (50). Although the mutations in FUS account for only about ~4% of FALS and rare cases of SALS (51), this mutation emphasizes the importance of RNA metabolism in ALS.
The presence of FUS-immunoreactive inclusions but not TDP-43 in protein aggregates are the pathological hallmark of FUS cases (27, 52). Using zebrafish, it has been shown that wild-type FUS rescues TARDBP knockdown but not vice versa therefore, indicating that FUS functions downstream of TDP-43 in the same RNA maturation pathway (53).

**Chromosome 9 open reading frame 72 (C9ORF72)**

The C9ORF72 gene is located on chromosome 9p21.2 and the gene gives rise to three transcripts of unknown function which are variant 1 (V1, NM_145005.6), variant 2 (V2, NM_0.18325.4) and variant 3 (V3, NM_001256054.2). Variant 2 and variant 3 encode for the long isoform (isoform a; C9ORF72-L) that consists of 481 amino acids, whereas variant 1 encodes the short isoform (isoform b; C9ORF72-S) that comprises of 222 amino acids. Genetic linkage analysis of a Scandinavian family with multiple family members affected with ALS and FTD had originally identified a locus on chromosome 9p21.3-p13.3 to be linked to ALS-FTD (54). Since then, hexanucleotide repeat expansions (HRE) in C9ORF72 have been identified as the most frequent genetic cause of familial ALS and familial frontotemporal degeneration (FTD) (23, 55). The pathogenic C9ORF72 HRE is found in the first intron of the variant 1 and variant 3 and is located within the predicted promoter region of variant 2 (Fig.3).

![C9ORF72 variants](image)

**Fig.3: The C9ORF72 hexanucleotide repeat expansions (HREs).** This schematic diagram shows the sites of C9ORF72 hexanucleotide repeat expansions, GGGGCC depicted as stars and the three annotated transcript variants; variant 1, variant 2 and variant 3.

The mutation in C9ORF72 has also been identified as being the most common mutation which is found in 38% of our Imperial College cohort, with similar findings in other European studies (23, 55-57). In a Taiwanese population, C9ORF72 HRE is still an important ALS-causing mutation with a frequency of 18%, indicating the global significance of this mutation (58). The number of hexanucleotide repeats in the intronic region of C9ORF72 varies from thirty to several hundred repeats (56). There have also

www.turkjphysiotherrehabil.org
been observed that patients with pure ALS, pure FTD or ALS-FTD may have 700-1600 repeats and the expansion size may be up to 10 kb in length, on the other hand, repeats found in controls are usually not greater than 30 repeats (23, 55). C9ORF72 encodes for a 481 amino acid (full length) protein that has an unknown function (12).

**ALS patients with C9ORF72 hexanucleotide repeat expansions (C9ORF72 ALS)** exhibit typical TDP-43 cytoplasmic inclusions in upper, lower, brainstem and spinal cord motor neurons however, p62-positive and C9ORF72 HRE-negative neuronal inclusions are found in cerebellum and hippocampus, that are specific to the C9ORF72-ALS patients (59, 60). These star-like neuronal inclusions are negative for TDP-43 and are composed of dipeptide repeat (DPR) proteins from unconventional repeat-associated non-ATG (RAN) translation of C9ORF72 expansion repeats (61-64).

### C9ORF72 transcripts

C9ORF72 mRNA is expressed in various brain regions and spinal cord with the highest expression observed in cerebellum (23, 55). C9ORF72 transcripts have been reported to be significantly decreased in many studies in different tissues and cell lines. C9ORF72 encodes for three different transcripts which in turn encode for two isoforms (isoforms a and b) and reduced level of mRNA for isoform a, the longer isoform has been observed in ALS cerebral cortex (DeJesus-Hernandez et al., 2011, Renton et al., 2011). Total C9ORF72 mRNA that consists of C9ORF72 isoform a and isoform b has been reported to be reduced in frontal cortex and cerebellum in frontotemporal degeneration (FTD) and ALS patients carrying C9ORF72 HRE (65-68). There have also been numerous reports of the reduction in the C9ORF72 isoform a transcript in the frontal cortex, motor cortex, cerebellum and cervical spinal cord of C9ORF72 HRE carriers (23, 66-70). The mRNA level of C9ORF72 isoform b has also been shown to be reduced in the frontal cortex and cerebellum of HRE carriers (65, 66).

Furthermore, inconsistencies have been observed in studies on C9ORF72 mRNA levels as shown in Table 4 and this could be due to variability in sample size, methodology, tissue type and RNA integrity. As shown also in Table 4, studies in brain tissues were performed in frontal cortex, cerebellum, motor cortex and there were only two studies performed in spinal cords samples. Also, the studies that were carried in cell lines utilised fibroblasts, iPSC and lymphoblastoid cell lines (LCLs) from ALS/FTD carriers. Therefore, our project is important because it utilised lumbar spinal cord from sporadic ALS cases and lymphoblastoid cell lines generated from familial ALS subjects. The previous studies were also mainly focusing on the mRNA levels of total C9ORF72 and C9ORF72 isoform a transcript. However, our project aimed to investigate the expressions of C9ORF72 isoforms a and b in lumbar spinal cords from SALS subjects, lymphoblastoid cell lines derived from C9ORF72 FALS individuals and in frontal and temporal...
cortex samples from frontotemporal dementia (FTD) subjects, with and without the presence of C9ORF72 HRE and Alzheimer's disease (AD) subjects.
Table 4: Summary of C9ORF72 mRNA expression in various tissues and cell lines published in literatures.

<table>
<thead>
<tr>
<th>REFERENCE</th>
<th>TRANSCRIPT</th>
<th>TISSUES STUDIED</th>
<th>NO. OF CONTROLS</th>
<th>NO. OF CASES</th>
<th>p-VALUE</th>
<th>REGULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>(23)</td>
<td>V2</td>
<td>Frontal cortex</td>
<td>Controls, n = 7</td>
<td>FTD-TDP with GGGGCC, n = 7</td>
<td>p&lt;0.01 **</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td>V2 and V3</td>
<td>Frontal cortex</td>
<td>Controls, n = 7</td>
<td>FTD-TDP with GGGGCC, n = 7</td>
<td>p&lt;0.05 *</td>
<td>Decreased</td>
</tr>
<tr>
<td>(65)</td>
<td>Total C9ORF72</td>
<td>Frontal cortex</td>
<td>Controls, n = 9</td>
<td>FTD with GGGGCC, n = 2</td>
<td>0.034 *</td>
<td>Decreased 50%</td>
</tr>
<tr>
<td></td>
<td>Isoform a (V2 and V3)</td>
<td>Frontal cortex</td>
<td>Controls, n = 9</td>
<td>FTD without GGGGCC, n = 2</td>
<td>0.034 *</td>
<td>Decreased 50%</td>
</tr>
<tr>
<td></td>
<td>Isoform b (V1)</td>
<td>Frontal cortex</td>
<td>Controls, n = 9</td>
<td>FTD without GGGGCC, n = 2</td>
<td>0.034 *</td>
<td>Decreased</td>
</tr>
<tr>
<td>(68)</td>
<td>Total C9ORF72</td>
<td>Frontal cortex</td>
<td>Controls, n = 5</td>
<td>FTD and ALS with GGGGCC, n = 10</td>
<td>0.014 *</td>
<td>Decreased</td>
</tr>
</tbody>
</table>

www.turkphysiother rehabil.org
<table>
<thead>
<tr>
<th>Isoform a (V2 and V3)</th>
<th>Frontal cortex</th>
<th>Controls, n = 5</th>
<th>FTD and ALS with GGGGCC, n = 10</th>
<th>0.089</th>
<th>Decreased</th>
</tr>
</thead>
<tbody>
<tr>
<td>Isoform b (V1)</td>
<td>Frontal cortex</td>
<td>Controls, n = 5</td>
<td>FTD and ALS with GGGGCC, n = 10</td>
<td>0.089</td>
<td>No change</td>
</tr>
<tr>
<td>Intron containing 1b (V2)</td>
<td>Frontal cortex</td>
<td>Controls, n = 5</td>
<td>FTD and ALS with GGGGCC, n = 8</td>
<td>7.9e-06</td>
<td>Decreased</td>
</tr>
</tbody>
</table>

(66) Taqman Assay

<table>
<thead>
<tr>
<th>Total C9ORF72</th>
<th>Cerebellum</th>
<th>Controls, n = 20</th>
<th>C9+ve, n = 52</th>
<th>2.61e-05</th>
<th>Decreased</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Controls, n = 20</td>
<td>C9-ve, n = 29</td>
<td>9.14e-06</td>
<td>Decreased</td>
</tr>
<tr>
<td>V2</td>
<td>Cerebellum</td>
<td>Controls, n = 20</td>
<td>C9+ve, n = 52</td>
<td>1.26e-06</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 20</td>
<td>C9-ve, n = 29</td>
<td>5.20e-09</td>
<td>Decreased</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total C9ORF72</th>
<th>Frontal cortex</th>
<th>Controls, n = 20</th>
<th>C9+ve, n = 52</th>
<th>0.02</th>
<th>Decreased</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Controls, n = 20 vs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C9-ve, n = 29</td>
<td></td>
<td>2.03e-07</td>
<td>Decreased</td>
</tr>
</tbody>
</table>

www.turkphysiotherrehabil.org
<table>
<thead>
<tr>
<th>V2</th>
<th>Frontal cortex</th>
<th>Controls, n = 20</th>
<th>C9+ve, n = 52</th>
<th>1.11e-05</th>
<th>Decreased</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Controls, n = 20 vs C9-ve, n = 31</td>
<td>-</td>
<td>8.53e-13</td>
<td>Decreased</td>
</tr>
<tr>
<td>V3</td>
<td>Cerebellum</td>
<td>Controls, n = 19</td>
<td>C9+ve, n = 54</td>
<td>-</td>
<td>No change</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 19 vs C9-ve, n = 29</td>
<td>-</td>
<td>-</td>
<td>No change</td>
</tr>
<tr>
<td>V3</td>
<td>Frontal cortex</td>
<td>Controls, n = 20</td>
<td>C9+ve, n = 52</td>
<td>-</td>
<td>No change</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 20 vs C9-ve, n = 29</td>
<td>-</td>
<td>-</td>
<td>No change</td>
</tr>
<tr>
<td>NanoString Technologies</td>
<td>Cerebellum</td>
<td>Controls, n = 19</td>
<td>C9+ve, n = 54</td>
<td>P-value</td>
<td>Significance</td>
</tr>
<tr>
<td>-------------------------</td>
<td>------------</td>
<td>-----------------</td>
<td>---------------</td>
<td>---------</td>
<td>--------------</td>
</tr>
<tr>
<td>Total C9ORF72</td>
<td></td>
<td></td>
<td></td>
<td>1.59e-07</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C9-ve, n = 29</td>
<td>-</td>
<td>0.003</td>
<td>Decreased</td>
</tr>
<tr>
<td>V2</td>
<td>Cerebellum</td>
<td>Controls, n = 19</td>
<td>C9+ve, n = 54</td>
<td>5.23e-10</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 19</td>
<td>C9-ve, n = 29</td>
<td>7.27e-07</td>
<td>Decreased</td>
</tr>
<tr>
<td>Isoform b (V1)</td>
<td>Cerebellum</td>
<td>Controls, n = 19</td>
<td>C9+ve, n = 54</td>
<td>9.04e-06</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 19</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C9-ve, n = 29</td>
<td>-</td>
<td>0.01</td>
<td>Decreased</td>
</tr>
<tr>
<td>Intron containing 1a (V1 and V3)</td>
<td>Cerebellum</td>
<td>Controls, n = 19</td>
<td>C9+ve, n = 54</td>
<td></td>
<td>-</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 19</td>
<td>C9-ve, n = 29</td>
<td>-</td>
<td>No change</td>
</tr>
<tr>
<td>Intron containing 1b (V2)</td>
<td>Cerebellum</td>
<td>Controls, n = 19</td>
<td>C9+ve, n = 54</td>
<td>-</td>
<td>No change</td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------</td>
<td>-----------------</td>
<td>----------------</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 19 vs C9-ve, n = 29</td>
<td>-</td>
<td>-</td>
<td>No change</td>
</tr>
<tr>
<td>V2</td>
<td>Frontal cortex</td>
<td>Controls, n = 20</td>
<td>C9+ve, n = 52</td>
<td>5.07e-10</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 20</td>
<td>C9-ve, n = 29</td>
<td>1.84e-08</td>
<td>Decreased</td>
</tr>
<tr>
<td>V1</td>
<td>Frontal cortex</td>
<td>Controls, n = 20</td>
<td>C9+ve, n = 52</td>
<td>0.001</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 20 vs C9-ve, n = 29</td>
<td>-</td>
<td>0.0001</td>
<td>Decreased</td>
</tr>
<tr>
<td>Intron containing 1a (V1 and V3)</td>
<td>Frontal cortex</td>
<td>Controls, n = 20</td>
<td>C9+ve, n = 52</td>
<td>0.003</td>
<td>Decreased</td>
</tr>
<tr>
<td>Intron containing 1b (V2)</td>
<td>Frontal cortex</td>
<td>Controls, n = 20</td>
<td>C9+ve, n = 52</td>
<td>-</td>
<td>No change</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------</td>
<td>-----------------</td>
<td>--------------</td>
<td>---</td>
<td>-----------</td>
</tr>
<tr>
<td>Controls, n = 19</td>
<td></td>
<td></td>
<td>-</td>
<td>-</td>
<td>No change</td>
</tr>
<tr>
<td>vs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C9-ve, n = 29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(71)</td>
<td>Total C9</td>
<td>Controls, n = 10</td>
<td>C9+ve ALS/FTD, n = 9</td>
<td>p&lt;0.05</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C9-ve ALS/FTD,</td>
<td>C9+ve ALS/FTD, n = 9</td>
<td>p&lt;0.005</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 9</td>
<td></td>
<td>***</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>C9-ve ALS/FTD,</td>
<td></td>
<td>p&lt;0.05</td>
<td>Increased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 9</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>vs</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controls, n = 10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoform a</td>
<td>Frontal Cortex</td>
<td>Controls, n = 10</td>
<td>C9+ve ALS/FTD, n = 9</td>
<td>p&lt;0.05</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>C9-ve ALS/FTD,</td>
<td>C9+ve ALS/FTD, n = 9</td>
<td>p&lt;0.05</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td></td>
<td>n = 9</td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>Total C9</td>
<td>Cerebellum</td>
<td>C9-ve ALS/FTD, n = 9 vs Controls, n = 10</td>
<td>Controls, n = 10</td>
<td>C9+ve ALS/FTD, n = 9</td>
<td>p&lt;0.01 **</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoform a</td>
<td>Cerebellum</td>
<td>C9-ve ALS/FTD, n = 9 vs Controls, n = 10</td>
<td>Controls, n = 10</td>
<td>C9+ve ALS/FTD, n = 9</td>
<td>p&lt;0.01 **</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Isoform</td>
<td>Region</td>
<td>Group 1</td>
<td>Group 2</td>
<td>p-value</td>
<td>Result</td>
</tr>
<tr>
<td>---------</td>
<td>--------</td>
<td>---------</td>
<td>---------</td>
<td>---------</td>
<td>--------</td>
</tr>
<tr>
<td>(69)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cerebellum</td>
<td>Controls, n = 4</td>
<td>ALS GGGGCC carriers, n = 6</td>
<td>p&lt;0.05</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td>Motor cortex</td>
<td>Controls, n = 10</td>
<td>ALS GGGGCC carriers, n= 6</td>
<td>p&lt; 0.001</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td>Cervical spinal cord</td>
<td>Controls, n = 4</td>
<td>ALS GGGGCC carriers, n= 6</td>
<td>p&lt; 0.05</td>
<td>Decreased</td>
</tr>
<tr>
<td>V2</td>
<td>Cerebellum</td>
<td>Controls, n = 4</td>
<td>ALS GGGGCC carriers, n = 6</td>
<td>p&lt;0.05</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td>Motor cortex</td>
<td>Controls, n = 10</td>
<td>ALS GGGGCC carriers, n= 6</td>
<td>p&lt; 0.001</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td>Cervical spinal cord</td>
<td>Controls, n = 4</td>
<td>ALS GGGGCC carriers, n= 6</td>
<td>p&lt; 0.05</td>
<td>Decreased</td>
</tr>
<tr>
<td>(70)</td>
<td>Isoform a (V2 and V3)</td>
<td>Motor cortex</td>
<td>Controls, n = 12</td>
<td>ALS GGGGCC carriers, n= 10</td>
<td>p&lt; 0.001 ***</td>
</tr>
<tr>
<td>------</td>
<td>----------------------</td>
<td>--------------</td>
<td>-----------------</td>
<td>-----------------------------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td>Spinal cord</td>
<td>Controls, n = 8</td>
<td>ALS GGGGCC carriers, n= 5</td>
<td>p&lt; 0.05 *</td>
<td>Decreased</td>
</tr>
</tbody>
</table>

| (67) | Total C9ORF72         | Frontal cortex | Controls, n = 8 | ALS/FTD GGGGCC carriers, n= 7 | 0.002 ** | Decreased |

<p>| Total C9ORF72 | Frontal cortex | Controls, n = 8 vs Sporadic FTD C9-ve, n= 5 | 0.03 * | Decreased |</p>
<table>
<thead>
<tr>
<th>REFERENCE</th>
<th>TRANSCRIPT</th>
<th>TISSUES STUDIED</th>
<th>NO. OF CONTROLS</th>
<th>NO. OF CASES</th>
<th>p-VALUE</th>
<th>REGULATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>(23)</td>
<td>Isoform b (V1)</td>
<td>LCLs</td>
<td>7</td>
<td>ALS GGGGCC carriers, n = 7</td>
<td>p&lt;0.01 **</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td>Intron containing 1a (V1 &amp; V3)</td>
<td>LCLs</td>
<td>7</td>
<td>ALS GGGGCC carriers, n = 7</td>
<td>p&lt;0.05 *</td>
<td>Decreased</td>
</tr>
<tr>
<td>(69)</td>
<td>Isoform b (V1)</td>
<td>Fibroblast</td>
<td>5</td>
<td>ALS GGGGCC carriers, n = 5</td>
<td>not significant</td>
<td>No Change</td>
</tr>
<tr>
<td></td>
<td>Isoform b (V1)</td>
<td>iPSC</td>
<td>3</td>
<td>ALS GGGGCC carriers, n = 5</td>
<td>p&lt;0.001 ***</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>Fibroblast</td>
<td>5</td>
<td>ALS GGGGCC carriers, n = 5</td>
<td>not significant</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td>V2</td>
<td>iPSC</td>
<td>3</td>
<td>ALS GGGGCC carriers, n = 5</td>
<td>p&lt;0.001 ***</td>
<td>Decreased</td>
</tr>
<tr>
<td>(70)</td>
<td>Isoform a (V2 and V3)</td>
<td>LCLs</td>
<td>5</td>
<td>ALS GGGGCC carriers, n = 6</td>
<td>p&lt;0.001 ***</td>
<td>Decreased</td>
</tr>
<tr>
<td>(67)</td>
<td>Total C9ORF72</td>
<td>LCLs</td>
<td>6</td>
<td>ALS/FTD GGGGCC carriers, n = 3</td>
<td>0.01</td>
<td>Decreased</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>-------</td>
<td>----</td>
<td>-------------------------------</td>
<td>------</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td>Isoform a (V2 and V3)</td>
<td>LCLs</td>
<td>6</td>
<td>ALS/FTD GGGGCC carriers, n = 3</td>
<td>0.04</td>
<td>Decreased</td>
</tr>
<tr>
<td>(72)</td>
<td>Total C9ORF72</td>
<td>LCLs</td>
<td>3</td>
<td>Sporadic ALS/FTD C9-ve, n = 5</td>
<td>0.04</td>
<td>Decreased</td>
</tr>
<tr>
<td></td>
<td>Isoform b (V1)</td>
<td>Neurons</td>
<td>3</td>
<td>ALS GGGGCC carriers, n = 3</td>
<td>P&lt;0.001***</td>
<td>Decreased</td>
</tr>
</tbody>
</table>

Table 4: Summary of C9ORF72 mRNA expression in various tissues and cell lines published in literatures. (Abbreviations: C9+ve, C9-positive; C9-ve, C9-negative; LCLs, lymphoblastoid cell lines; iPSC, induced pluripotent stem cells).
Clinical phenotypes associated with C9ORF72 hexanucleotide repeat expansions

C9ORF72 HRE give rise to a more aggressive form of ALS than other mutation, where patients with HRE have an earlier age at onset, about ~2.5 years earlier than patients without the repeat expansions and are more likely to have a positive family history of dementia (56). ALS patients carrying C9ORF72 HREs also have shorter survival, about 6 months shorter than the patients without the repeat (73). Initial studies have suggested that C9ORF72-ALS patients are also more likely to show bulbar-onset symptoms than limb symptoms compared to patients without C9ORF72 HRE (74-76) but this has not been confirmed in subsequent cohort studies (73, 77). ALS patients carrying C9ORF72 HREs are indistinguishable from classic ALS patients but nearly 50% of the C9ORF72-positive ALS cases develop cognitive or behavioural impairment or both, compared to those without C9ORF72 HRE where these features are less common (78, 79). Progressive muscular atrophy or primary lateral sclerosis is not common in ALS patients carrying HRE (56).

FTD patients carrying C9ORF72 HREs are commonly present with behavioural variant FTD (bvFTD) in which the diagnostic criteria include apathy, disinhibition and socially inappropriate behaviour, loss of empathy and perseverative and stereotyped or obsessive-compulsive behaviour. Furthermore, 28-56% of bvFTD patients carrying C9 HREs have hallucinations and delusions compared with 4-18% of patients without HRE (61, 80-87).
The mechanisms of neurodegeneration in C9ORF72

The mechanism of neurodegeneration in C9ORF72 HRE is still unclear but it is proposed to occur through loss-of-function, a gain-of-function or perhaps both due to abnormal protein formation or abnormal RNA toxicity (88).

Loss of C9ORF72 function

Reduced C9ORF72 mRNA levels in brain and lymphoblasts of C9ORF72 repeat expansion carriers have been reported, suggesting loss-of-function by C9ORF72 haploinsufficiency may contribute to the pathogenicity of the C9 ALS/FTD (23, 55, 65). Consequently, C9ORF72 mRNA levels were decreased in the frontal cortex, cerebellum, hippocampus and spinal cord of C9 HRE carriers in comparison to non-HRE subjects and controls and this was also observed in the C9ORF72 protein levels in the frontal and cortex samples (68). Translations of C9ORF72 mRNA gives rise to two C9ORF72 protein isoforms which are long isoform (C9ORF72-L) and short isoform (C9ORF72-S). Western blotting analysis has also revealed in the brain, C9ORF72-L protein level was reduced whereas the C9ORF72-S protein level was increased. Further, in the spinal cord motor neurons, C9ORF72-S protein expression was reduced on the nuclear membrane of C9ORF72-positive ALS compared to controls (89). Interestingly, no nuclear membrane staining of C9ORF72-L was observed in the spinal cord motor neurons (89). C9ORF72-L protein expression was also reported to be 80% decreased in the cerebellum of C9ORF72 HRE carriers compared to healthy control (90).

Several mechanisms might lead to loss-of-function in C9ORF72. Hypermethylation of a CpG island upstream of the repeats and trimethylation of lysine residues on H3 and H4 histones, which bind GGGGCC repeats may lead to transcriptional silencing of C9ORF72 (71, 91, 92). Hypermethylation of C9ORF72 HRE occurs when the number of repeats is greater than 90 (92) and it has been demonstrated in human kidney and neuroblastoma cell lines that C9ORF72 transcription is downregulated in cells with a high number of repeats and methylations (93). C9ORF72 is structurally homologous to Differentially Expressed in Normal and Neoplasia (DENN) proteins that act as Guanine nucleotide exchange factors (GEFs) to activate RAB GTPases and may therefore, regulate membrane trafficking (94, 95). C9ORF72 has also been shown to interact with Rab1a and the Unc-51-like kinase 1 (ULK1) autophagy initiation complex and C9ORF72 regulates the initiation of autophagy by regulating the translocation of ULK1 autophagy initiation complex to the phagophore (96). Loss of this function leads to the accumulation of p62 in HeLa and primary cortical neurons (96) that mimic the p62 pathology of C9 ALS/FTD patients with characteristic p62-positive, neuronal cytoplasmic and...
intranuclear inclusions in the cerebellum and hippocampus (73, 97, 98). It has also been reported that knockdown of C9ORF72 protein in neuronal cell lines causes endocytosis and autophagy defects (99) and knockdown of zC9ORF72, the C9ORF72 orthologue in zebrafish resulted in axonopathy and motor deficits (67) whereas knockout of the C9ORF72 orthologue, afa-1 in Caenorhabditis elegans model resulted in motility defects leading to paralysis (100). However, motor neuron degeneration which is characterised of ALS (101) and motor deficits have not been observed in C9ORF72 knockout mouse model (102). These discrepancies may be caused by the high degree of homology between human C9ORF72 and mouse orthologue (98%) compared with its homology to zebrafish (76%) or Caenorhabditis elegans (58%) (67, 100). Further, C9ORF72 knockout mice showed mild motor deficits and altered autoimmune response including increased activation of T cells and plasma cells, increased autoantibodies and evidence of progressive glomerulonephropathy thus implicating a role of C9ORF72 protein in the regulation of immune homeostasis (103).

**Toxic Gain-of-Function through the generation of Aberrant RNA Foci**

It has also been suggested that C9ORF72 repeat expansions may result in a gain-of-toxic function mechanism through RNA-mediated toxicity by forming RNA foci that may sequester RNA binding proteins. The RNA binding proteins (RBPs) that have been identified to be sequestered by the repeat RNA foci are SRSF2, hnRNP H1/F, ALYREF, hnRNPA3, hnRNPA1, hnRNP-H, nucleolin, Pur-α, ASF/SF2, ADARB2 and RanGAP1 (62, 69, 70, 104-108). There have been observed that sense and antisense RNA foci are present in peripheral tissues namely skin derived-fibroblasts, lymphoblasts, spinal cord, hippocampus, cerebellum and frontal cortex of ALS patients carrying HRE (23, 67, 102, 109-111). Studies have also reported that RNA foci and DPR proteins are detected in the C9 HRE BAC-transgenic mice model but these mice model displayed no or only mild neurodegeneration (112-115). Therefore, it might be possible that C9ORF72 haploinsufficiency may result in a modified C9ORF72 HRE toxic gain-of-function such as reduced autophagy due to C9ORF72 loss-of-function which results in DPR proteins accumulation and toxicity (96).

Aberrant HRE-containing mRNA can also form two structures, G-quadruplexes (G-Q) which is a highly stable secondary structure formed by the stacking of planar tetrads of four nonsequential Guanosine residues (G-quartets) and hybrids with HRE-containing DNA to form R-loops (70, 116-118). A recent study has identified three small molecules (DB 1246, DB 1247 and DB 1273) that are non-conjugated aromatic diamines which have similar structures and are able to bind and stabilise the HRE quadruplexes that results in decreased nuclear RNA foci and DPR formation in iPSC derived neurons from C9ORF2 ALS patient (119). One of the small molecules (DB 1273) has also been observed to
decrease repeat associated non-AUG (RAN) translation and dipeptide repeat (DPR) proteins formations and also improve the survival of GGGGCC repeat-expressing *D. melanogaster* (119). C-rich sense and antisense strands of DNA containing HRE can assemble to form i-motifs and hairpin structures (120). An earlier study has also reported that small molecules were able to bind to the hairpin conformation of HREs and significantly reduce RNA foci and DPR proteins (121).

**Toxic Gain-of-Function due to Accumulation of Dipeptide Repeat (DPR) Proteins**

The translation of repeat associated non-AUG (RAN) of C9ORF72 may result in aberrant dipeptide repeat (DPR) proteins composed of glycine-alanine (GA), glycine-proline (GP), glycine-arginine (GR) in the sense frames and glycine-proline (GP), alanine-proline (AP) and proline-arginine (PA) in the antisense frames (122) with GP is the most abundant in C9 ALS/FTD (111). The accumulations of DPR proteins are prone to aggregation and form neuronal inclusions in the brain of C9 ALS/FTD patients that colocalize with p62 but not TDP-43 (62, 63, 109, 111, 123, 124). DPR-based gain-of-function was demonstrated in cultured neurons, where expression of poly-GA peptides generates cytoplasmic aggregate inclusions, reduces dendritic branching, induces endoplasmic reticulum stress and enhances apoptosis (125, 126). Additionally, the arginine-rich DPR proteins, glycine-arginine (GR) and proline-arginine (PR) are retained in cells, resulting in transcriptional dysregulation therefore, causing cell death (127), eye degeneration and decreased survival in *Drosophila* (128). DPR proteins composed of arginine especially poly-PR are most harmful because arginine-rich DPR proteins have been reported to interrupt nucleocytoplasmic transport (129) and RNA processing (127, 129) which result in translation arrest (130, 131), nucleolar stress (132), impairment of ubiquitin proteasome system (133) and disrupt the formation of stress granules (134). DPR proteins consisted of poly-GA can also disturb the ubiquitin proteasome system and cause endoplasmic reticulum stress (126).

In summary, it is evident that hexanucleotide repeat expansions in C9ORF72 result in both loss-of-function and gain-of-function mechanisms, however, the relative contribution of these mechanisms in ALS pathogenesis still needs to be investigated.
Other Rare Genes Causing ALS

Ubiquilin 2 (UBQLN2), sequestosome 1 (SQSTM1) and valosin-containing protein (VCP)

Recently, mutations in proteins involved in proteostasis namely ubiquilin 2; UBQLN2 (135), Sequestosome 1; SQSTM1 (29) and Valosin-containing protein; VCP (28) have also been found to cause FALS. UBQLN2 is an X-linked gene that encodes for ubiquilin 2 and is a rare mutation in FALS (30, 135). UBQLN2 binds the ubiquitinilated proteins through its UBA domain and to the proteasome through its UBL domain for protein degradation through ubiquitin-proteasomal system (136). SQSTM1 is the gene encoding for the p62 protein, that has been shown to coexist in FALS cases with C9ORF72 mutations (29). p62 is a stress-inducible intracellular protein and it interacts with microtubule associated protein, LC3 and is then incorporated into the autophagosome for degradation by autophagy (137). VCP encodes for valosin-containing protein (p97) and the protein is important in cell cycle progression, proteosomal degradation and membrane dynamics and is also essential for growth in mice and other model organisms (138). VCP mutations are the cause of approximately ~1% of FALS (28, 139).

Optineurin (OPTN), angiogenin (ANG) and vesicle-associated membrane protein B (VAPB)

A mutation in the Optineurin gene (OPTN) has been identified in the Japanese population and is very rare in people of European descent (140, 141). However, the Optineurin gene (OPTN) has been identified as a genetic factor in GWAS studies of Paget’s disease of bone together with SQSTM1 and VCP and this provides an explanation about the overlap of clinical presentations of ALS and Paget’s disease of bone (142-144). Approximately ~1% of FALS cases are due to a mutation in Angiogenin (ANG) gene, which encodes angiogenin and ribonuclease 5 (4). The VAPB gene on chromosome 20 encodes for vesicle-associated membrane protein B (VAPB) (145) and it reduces the accumulation of unfolded proteins through the endoplasmic reticulum (ER) unfolded protein response (UPR) (146). VAPB levels were found to be significantly reduced in the spinal cord of ALS individuals (147) and in the SOD1 mouse model of ALS (148). A VAPB mutation has also been reported in one FALS patient carrying C9 HRE (149) and this suggests that ALS has an oligogenic aetiology, as it has previously been shown that 5 out of 97 Dutch FALS families have multiple mutations in ALS-associated genes (150).
Turbin alpha 4a (TUBA4A)

Tubulin alpha 4a (TUBA4A) gene on chromosome 2q35 has been identified as a novel causal gene in ALS and FTD and 10 missense, 1 nonsense and 1 splice donor site mutations have been reported in SALS and FALS cases with some cases presenting with FTD (151-153). However, a large exome sequencing study consisting of 3000 ALS patients and 600 controls did not observe association for TUBA4A (154). More recently, a mutation screening in a Belgian population consisting of 459 FTD cases, 28 FTD-ALS cases, 429 ALS cases and 703 controls has reported a novel TUBA4A frameshift mutation (p.Arg64Glyfs*90) resulting in a truncated protein with loss of the C-terminal half of the last exon 4 in a FTD patient (152). An earlier study has also described a nonsense mutation (p.Trp407*) resulting in the loss of 41 amino acids at the C-terminal that is essential for the polymerisation with the B-tubulins to incorporate microtubule cytoskeleton (151). Additionally, a patient specific missense mutation (p.Thr381Met) has been observed in a FALS patient and this was also present in her brother affected with ALS and memory problems (152) and interestingly, both of the FALS patients are also C9ORF72 HRE carriers (152).

TANK-binding kinase 1 (TBK1)

Mutations in TANK-binding kinase 1 (TBK1), also known as T2K or NAK gene have also been linked to ALS. A large-scale exome sequencing study involved 2869 ALS patients and 6405 controls has identified TBK1 as a novel ALS causal gene (154). TBK1 is a ubiquitously expressed kinase and is involved in innate immunity as an inducer of type-1 interferons. TBK1 is also important in autophagy as it is involved in the phosphorylation of autophagy adaptors for examples, p62, OPTN and NDP52 to enhance bindings to ubiquitinated cargos and LC3-II (155, 156). TBK1 also functions in the insulin signalling pathway (157) and innate immunity against bacterial and viral infections (158, 159). TBK1 was originally implicated in primary open glaucoma and normal tension glaucoma due to gain-of-function through increased copy number of TBK1 (160, 161) and is involved in herpes simplex encephalitis in childhood possibly due to reduced activity in TLR3 mediated immunity (162). Currently, TBK1 variants have been reported in ALS patients from many different populations worldwide for examples, in a Swedish population, Australian population of Chinese origin, Taiwanese, Chinese, Sardinian and Belgian populations (163-168). A study in a French population has observed that TBK1 mutation is more common in FTD patients comorbid with ALS (10.8%) compared to patients with ALS only (0.5%) (169).

Mutations reported in TBK1 are nonsense, frameshift, missense and deletions all are which contribute differently to ALS pathogenesis (170). Nonsense and frameshift mutations may lead to decreased
mRNA and protein expressions therefore, causing haploinsufficiency (163, 171). In vitro study has reported loss of expression due to loss-of-function Tbk1 mutant alleles or loss of interaction of the C-terminal TBK1 coiled-coil domain (CCD2) mutants with the TBK1 adaptor protein, optineurin (OPTN) resulted in impaired autophagy that can contribute to cytoplasmic aggregates (169). A study has reported that TBK1 colocalizes with Rab8b and LC3 on autophagosomes therefore suggesting that TBK1 is involved in autophagosome maturation (155).

NIMA (never in mitosis gene a)-related kinase 1 family (NEK1)

Further, whole exome sequencing studies have reported a significant enrichment of NEK1 heterozygous loss-of-function (LOF) variants in ALS (154, 172, 173). NEK1 is a serine/threonine kinase of the highly conserved NIMA (never in mitosis gene a)-related kinase 1 family. NEK1 is important in cell-cycle regulation, ciliogenesis, mitochondrial membrane permeability and DNA damage repair (174-176). NEK1 has been observed to localize at the DNA foci after γ-irradiation to facilitate the DNA damage response (DDR) (177, 178). NEK1 plays a key role in the activation of cell cycle kinases CHK1 and CHK2 (checkpoint 1/2) that are essential for proper arrest at G1/S, S- or G2/M-phase and DNA damage response (DDR) (177) which if unrepaired may lead to chromosomal breaks and accumulation of genomic rearrangements (179, 180). NEK1 also binds to C21ORF2, an ALS risk factor found through genome-wide association studies (GWAS) (154, 173, 181, 182) and the interactions between these two proteins are essential for DNA damage repair (183).

DNA damage has been implicated with fatal neurodegeneration and ALS progression (184-187). Currently, 33 NEK1 LOF variants have been reported including 10 variants that were also observed in unaffected controls (154, 172, 173, 188). A recent study in a Belgian population has observed 17 unique NEK1 variants with 7 variants (2 LOF and 5 missense) present in 13 from 278 patients (4.7%) and 14 variants (1 LOF, 12 missense and 1 amino acid deletion) in 20 from 609 control individuals (189). Interestingly, in the study by Nguyen et al. (189), they confirmed two FALS siblings were present with a NEK1 p.Ser1036* nonsense variant and the siblings also carried a TUBA4A p.Trh381Met variant and C9ORF72 repeat expansions which were described in an earlier study (65). This implies that the accumulation of mutations with variable penetrance may contribute to ALS disease manifestation (189).
Kinesin family member 5A (KIF5A)

Additionally, a large-scale GWAS involving 20,806 ALS and 59,804 controls has identified 1 missense mutation within kinesin family member 5A (KIF5A) that is a significant risk for ALS and 3 loss-of-function (LOF) variants within KIF5A (190). The missense mutation within KIF5A results in a p.Pro986Leu (rs113247976) coding change within KIF5A is a relatively common but low penetrance risk allele for ALS (190). The study also showed that LOF variants in ALS patients are clustered at the 5’ and 3’ splice junctions of exon 27 and in the C-terminal cargo-binding region of KIF5A (amino acids 907–1032) which results in disrupted C-terminal peptide sequence (190). These LOF variants are rare but high penetrance ALS risk factors (190). Interestingly, the study by Nicolas et al. (190) has reported that ALS subjects with KIF5A LOF mutations showed an early age of onset (46.5 years) compared to the age of onset reported for ALS in epidemiological studies (65.2 years) (191). However, the study (190) has also observed that they displayed a much longer median survival time, approximately 10 years compared to typical ALS patients that have 20-36 months survival (191).

CONCLUSION

In summary, studying the ALS causing genes provide insight into the pathophysiology of both familial and sporadic ALS. Aberrant RNA metabolism, dysfunctional nucleocytoplasmic transport and impaired protein homeostasis are some of the pathways implicated in the pathogenesis of ALS. Currently, there are limited effective drugs available for the treatment of ALS and a further understanding of the mechanisms leading to the pathogenesis of the disease would, therefore, provide insight on therapeutic targets for ALS intervention.

DECLARATIONS

Acknowledgements

This work is dedicated to the memory of Jackie de Belleroche, an intelligent woman with passion and dedication who has contributed to scientific progress in the study of ALS. She was a mentor to all of us in her lab and she will always be. Special thanks to Alex Morris, Midhat Salman and Luigi Montibeller from the Neurogenetics Group for their support and to Majlis Amanah Rakyat (MARA) Malaysia for the PhD financial assistance.
Contribution
IB wrote the manuscript.

Consent for publication
Not applicable.

Competing interest
The author declare that she has no competing interest.

REFERENCES


