Structural Variants May Be a Source of Missing Heritability in sALS

Frances Theunissen[1,2], Loren L. Flynn[1,3,4], Ryan S. Anderton[1,3,5], Frank Mastaglia[1,3], Julia Pytte[1,2], Leanne Jiang[1,6], Stuart Hodgetts[1,2], Daniel K. Burns[7], Ann Saunders[7], Sue Fletcher[3,4], Steve D. Wilton[1,3,4] and Patrick Anthony Akkari[1,3,4]*

[1] Perron Institute for Neurological and Translational Science, Nedlands, WA, Australia, [2] School of Human Sciences, University of Western Australia, Nedlands, WA, Australia, [3] Centre for Neuromuscular and Neurological Disorders, University of Western Australia, Nedlands, WA, Australia, [4] Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, WA, Australia, [5] School of Health Sciences, Institute for Health Research, University of Notre Dame Australia, Fremantle, WA, Australia, [6] School of Biological Sciences, University of Western Australia, Nedlands, WA, Australia, [7] Zinfandel Pharmaceuticals, Chapel Hill, NC, United States

The underlying genetic and molecular mechanisms that drive amyotrophic lateral sclerosis (ALS) remain poorly understood. Structural variants within the genome can play a significant role in neurodegenerative disease risk, such as the repeat expansion in *C9orf72* and the tri-nucleotide repeat in *ATXN2,* both of which are associated with familial and sporadic ALS. Many such structural variants reside in uncharacterized regions of the human genome, and have been under studied. Therefore, characterization of structural variants located in and around genes associated with ALS could provide insight into disease pathogenesis, and lead to the discovery of highly informative genetic tools for stratification in clinical trials. Such genomic variants may provide a deeper understanding of how gene expression can affect disease etiology, disease severity and trajectory, patient response to treatment, and may hold the key to understanding the genetics of sporadic ALS. This article outlines the current understanding of amyotrophic lateral sclerosis genetics and how structural variations may underpin some of the missing heritability of this disease.

Keywords: amyotrophic lateral sclerosis, structural variant, genetic marker, missing heritability, clinical trial stratification

AMYOTROPHIC LATERAL SCLEROSIS; CLINICAL PHENOTYPES AND HERITABILITY

Amyotrophic lateral sclerosis (ALS) is a devastating progressive neurodegenerative disorder characterized by the loss of lower and upper motor neurons, resulting in paralysis of the limb, bulbar and respiratory muscles, and is typically fatal within 3–5 years from disease onset (Robberecht and Philips, 2013). The disease incidence is ∼2 per 100,000 population and it is projected that rates of ALS will increase from ∼222,000 worldwide in 2015 to ∼376,000 in 2040 (Arthur et al., 2016). In addition to the growing financial burden of this disease, there is a significant burden on the patients' families and carers. As such, it is critical to improve our understanding of how genetic mechanisms may contribute to the pathogenesis of this devastating disease. Investigation of short structural variations (SVs) in known ALS genes has multiple potential objectives, and may help to uncover some of the missing heritability in sporadic ALS. Characterization of short SVs may inform

the discovery of novel disease mechanisms and therapeutic targets, and be useful for stratification of patient subgroups in clinical trials.

ALS is a heterogeneous disease, with variable clinical presentation between patients, and is characterized by progressive motor deficits that evolve over weeks or months, eventually affecting most voluntary muscles in the body (Talman et al., 2016). The heterogeneity of clinical presentation and diverse rates of progression make the disease challenging to diagnose, and there is currently no definitive diagnostic test. As a result, it is usually characterized on the basis of the site and pattern of onset, and the degree of involvement of upper and lower motor neurons, and can be classified into the following categories: (i) progressive bulbar palsy (difficulty swallowing and speech disturbances); (ii) limb-onset ALS; (iii) progressive muscular atrophy (involving only lower motor neurons); and (iv) upper motor neuron predominant ALS (Kinsley and Siddique, 2015). Most commonly, individuals will present with asymmetrical focal weakness of the extremities (poor hand grip, foot drop, stumbling) or bulbar characteristics (dysarthria and dysphagia). Other typical symptoms include muscle fasciculation, cramps and hyperreflexia in regions of atrophy, without accompanying sensory disturbances (Kinsley and Siddique, 2015). Importantly, the different clinical phenotypes exhibit differing rates of progression, with the bulbar-onset form having the most rapid progression and shortest survival time (Okamoto et al., 1993). In addition, there is considerable variability between cases within the same diagnostic category. Given the degree of heterogeneity, it remains to be determined whether the different clinical phenotypes all represent variations of the same disease, or whether there is also heterogeneity in the underlying genetic and molecular disease determinants. At present, there is a lack of specific genetic or other biomarkers for the different disease subtypes, or indicators of disease trajectory in well-documented patient cohorts.

Approximately 10% of ALS cases are classed as familial (fALS), while the remainder, with no prior family involvement, are classified as having the sporadic form of the disease (sALS) (Renton et al., 2014). Since the landmark discovery of the first mutation in the superoxide dismutase 1 gene (SOD1) in fALS in the early 1990s (Rosen et al., 1993), there has since been significant progress in understanding of the genetics of the familial disease, with approximately 70% of the genetic mutations that contribute to fALS having been identified (Cook and Petrucelli, 2019). However the genetic underpinnings of sporadic ALS (sALS) remain a formidable challenge (Al-Chalabi et al., 2012; Renton et al., 2014). Comparatively, few mutations have been described for sALS, despite the application of high throughput genetic analysis methods (Nguyen et al., 2018). These methodologies have failed to identify disease-associated genetic variations in the majority of sALS patients, highlighting the complexity and genetic heterogeneity contributing to this disease phenotype. Approximately 10% of sALS cases can be explained by mutations in 25 known ALS-linked genes, with the remaining 90% of cases as yet having an undetermined genetic contributor (Andersen and Al-Chalabi, 2011; Renton et al., 2014; Dharmadasa et al., 2017). At a glance these data may imply that

the genetic contributions to sALS are minor, however, heritability estimates and twin studies suggest a genetic contribution of up to 65% (Al-Chalabi et al., 2010; Al-Chalabi and Visscher, 2014). It is therefore likely that additional genetic contributors to sALS risk remain to be identified. The situation may be analogous to that for Alzheimer's disease where the *APOE* ε4 (Apolipoprotein E) risk allele has a frequency of only 14% and does not fully explain the Alzheimer's disease age-of-onset risk (Corder et al., 1993; Roses et al., 2016). However, after discovery of the structural variant (rs10524523) in the neighboring *TOMM40* (Translocase of outer mitochondrial membrane 40) gene, age of onset risk could now be assessed for >90% of the at risk population (Roses et al., 2010).

Currently there are only two therapeutics available for the treatment of ALS, Riluzole (approved in Australia, United States, and Europe) and Edaravone (approved in Japan, South Korea, and United States) (Rothstein, 2017) that impact excitotoxicity (Doble, 1996) or anti-oxidant pathways (Rothstein, 2017), respectively and may slow disease progression for a relatively short period of time. However, these drugs are only effective in some patients and there is currently no way to determine those most likely to respond to the drugs. For the patients that do show a response, life is only prolonged by approximately 3 months. Further understanding of ALS pathogenesis should inform the development of more effective therapies and help identify patients likely to respond to specific treatments.

GENETIC CHARACTERIZATION AND GENOME-WIDE ASSOCIATION STUDIES

Genome wide association studies analysing single nucleotide polymorphisms (SNP) and whole exome sequencing studies have provided a wealth of information relating to common variants associated with a range of diseases. Despite this, such approaches have generally identified genes that are either inherited in fALS, those that are weakly associated with sALS, or mutations associated with rare forms of the disease (Renton et al., 2014). Some of the ALS genes identified by these techniques include *UNC13A, C9orf72, DPP6, ELP3, KIFAP3, TBK1, CHCHD10, TUBA4A, CCNF, MATR3, NEK1, C21orf2, ANXA11,* and *TIA1* (see **Table 1**; Andersen and Al-Chalabi, 2011; Nguyen et al., 2018). A major issue with utilizing these technologies is that by testing thousands of SNPs across the genome, low effect size associations are generated for numerous gene regions, inevitably leading to very high thresholds of significance for potentially weak genetic effects (Roses, 2016). This can lead to false positive associations or a lack of reproducibility between different populations that have rarely translated into tools for clinical trial patient stratifications, with the exception of *UNC13A* (see below). The amount of information these techniques can provide regarding complex disease and the functional outcomes of SNPs are limited. Particularly, these methods have not been able to account for the variation in age-of-onset and progression in ALS patients (Al-Chalabi and Hardiman, 2013), and fail to explain the missing heritability of the disease (Al-Chalabi et al., 2017; Mejzini et al., 2019). Whole genome sequencing can

TABLE 1 | This table lists published genes that have been associated with ALS and highlights the discovery method for each gene as well as the gene function.

ALS genes	Discovery method	Gene function	References	Number of predicted short SVs
C9orf72	GWAS	RNA metabolism	Morita et al., 2006	31
TARDP	Candidate gene linkage	RNA metabolism	Gitcho et al., 2008	63
FUS	Candidate gene linkage	RNA metabolism	Vance et al., 2009	45
MATR3	WES	RNA metabolism	Johnson et al., 2014	60
TIA1	WES	RNA metabolism	Cirulli et al., 2015	83
HNRNPA1	Linkage WES	RNA metabolism	Kim et al., 2013	23
HNRNPA2/B1	Linkage WES	RNA metabolism	Kim et al., 2013	24
EWSR1	Candidate gene	RNA metabolism	Couthouis et al., 2012	66
TAF15	Candidate gene	RNA metabolism	Ticozzi et al., 2011	35
ANG	Candidate gene	RNA metabolism	Greenway et al., 2006	25
SMN1	Candidate gene	Interaction with RNA binding proteins	Corcia et al., 2002	68
ELP3	GWAS	Transcript elongation	Simpson et al., 2008	62
SETX	Linkage	DNA/RNA processing	Chen et al., 2004	101
SPG11	Linkage	DNA damage	Orlacchio et al., 2010	110
APEX1	Candidate gene	Endonuclease	Greenway et al., 2004	13
UBQLN2	Candidate gene linkage	Protein quality control	Deng et al., 2011	9
VCP	Candidate gene	Protein quality control	Johnson et al., 2010	37
OPTN	Homozygosity mapping	Protein quality control	Maruyama et al., 2010	54
VAPB	Linkage	Protein quality control	Nishimura et al., 2004	64
TBK1	WES	Protein quality control	Freischmidt et al., 2015	59
SQSTM1	Candidate gene	Protein quality control	Fecto et al., 2011	38
CCNF	Genome wide linkage	Protein quality control	Williams et al., 2016	48
PFN1	Linkage WES	Cytoskeletal and trafficking	Wu et al., 2012	21
TUBA4A	WES	Cytoskeletal and trafficking	Smith et al., 2014	24
KIF5A	GWAS	Cytoskeletal and trafficking	Nicolas et al., 2018	52
ANXA11	WES	Cytoskeletal and trafficking	Smith et al., 2017	41
NEFH	Candidate gene	Cytoskeletal and trafficking	Figlewicz et al., 1993	43
DCTN1	Candidate gene	Cytoskeletal and trafficking	Münch et al., 2004	18
PRPH	Candidate gene	Cytoskeletal protein	Leung et al., 2004	50
FIG4	Candidate gene	Cytoskeletal organization and vesicle trafficking	Chow et al., 2009	66
CFAP410	GWAS	Cytoskeletal and DNA damage response	Van Rheenen et al., 2016	15
KIFAP3	GWAS	Kinesin associated protein	Landers et al., 2009	55
ALS2	Linkage	Endosomal dynamics	Hadano et al., 2001	69
SIGMAR1	Homozygosity mapping	Endoplasmic reticulum chaperone	Al-Saif et al., 2011	21
SOD1	Linkage	Mitochondrial dysfunction and oxidative stress	Rosen et al., 1993	30
CHCHD10	Candidate gene	Mitochondrial dysfunction and oxidative stress	Bannwarth et al., 2014	32
NEK1	WES	Mitochondrial dysfunction and oxidative stress	Kenna et al., 2016	99
ATXN2	Candidate gene	Endocytosis, cell survival	Elden et al., 2010	186
GRN	Candidate gene	Cell growth regulator	Schymick et al., 2007	34
UNC13A	GWAS	Neurotransmitter release	Van Es et al., 2009	103
PLCD1	GWAS	Signal transduction	Staats et al., 2013	21
CHMP2B	Candidate gene	Recycling of cell receptors	Parkinson et al., 2006	22
ITPR2	GWAS	Receptor	Van Es et al., 2007	204
ARHGEF28	Candidate gene	Nucleotide exchange factor	Droppelmann et al., 2013	138
DAO	Candidate gene	Potential detoxifying agent	Mitchell et al., 2010	52
DPP6	GWAS	Modifies calcium gated channels	Van Es et al., 2008	235
VEGFA	Candidate gene	Angiogenesis, migration of endothelial cells	Lambrechts et al., 2003	15
HFE	Candidate gene	Iron absorption	Wang et al., 2004	31
PON1	Candidate gene	Organophosphate hydrolysis	Slowik et al., 2006	24

Using the method described by Saul et al. (2016) we have predicted the number of short structural variants in each gene that may warrant further investigation. This does not include insertion/deletions or SNPs.

examine the entirety of the genome to better capture larger scale variations, as opposed to single nucleotide changes, however, these methods come with their own technical limitations; high throughput short-read sequencing technologies are unable to accurately capture these regions due to amplification stutter, and misaligning of the short-read sequences, often misrepresenting their true variability (Cameron et al., 2019). Recently, it was demonstrated that variable areas of the genome can camouflage each other, particularly where short-read DNA sequences map equally well to different loci, resulting in poor SV characterization (Ebbert et al., 2019). To date, these techniques have only accounted for the genetic cause of approximately 10% of sALS cases (Cook and Petrucelli, 2019). Therefore, it is essential to rethink the approach and acknowledge the limitations of these technologies when interrogating the genome. Specifically, there are vast regions of genetic variability yet to be uncovered in non-coding regions, which might have significant implications in the context of complex disease.

STRUCTURAL VARIANTS

Structural variants (SVs) are defined as insertions, deletions, inversions and microsatellites that can be repeated hundreds of times. SVs predominantly occur in non-coding regions of the genome and often do not change the composition of the mature protein (Roses et al., 2016). Despite this, changes in the size and composition of SVs can have a significant impact on the regulatory elements that modulate gene expression (Chiang et al., 2017). Therefore, SVs can potentially provide a deeper understanding of how gene expression in complex genetic disease can affect disease etiology, duration, progression and patient outcomes (Feuk et al., 2006). SVs have been implicated in many complex diseases including retinitis pigmentosa (*MSR1*) (Rose et al., 2016), Alzheimer's (*TOMM40*) (Lyall et al., 2013), frontotemporal dementia (*C9orf72*) (DeJesus-Hernandez et al., 2011; Renton et al., 2011), and other neurodegenerative diseases (Beck et al., 2013).

The ability of SVs to alter gene expression is likely dependent on their location within and around the gene or intergenic region, with their effects occurring via several mechanisms including, influencing the binding of regulatory elements that determine transcription, mRNA splicing and processing, genome folding and higher order structure, and translation (Roses et al., 2016). This may differentiate mechanisms of disease pathogenesis, including risk of disease, risk for a specific phenotype, symptom presentation, disease course and response to treatment, between individuals (**Figure 1A**). Due to the variable nature, as well as the repeat structure of SVs, many remain poorly characterized by analysis platforms such as next generation sequencing (Cameron et al., 2019; Ebbert et al., 2019).
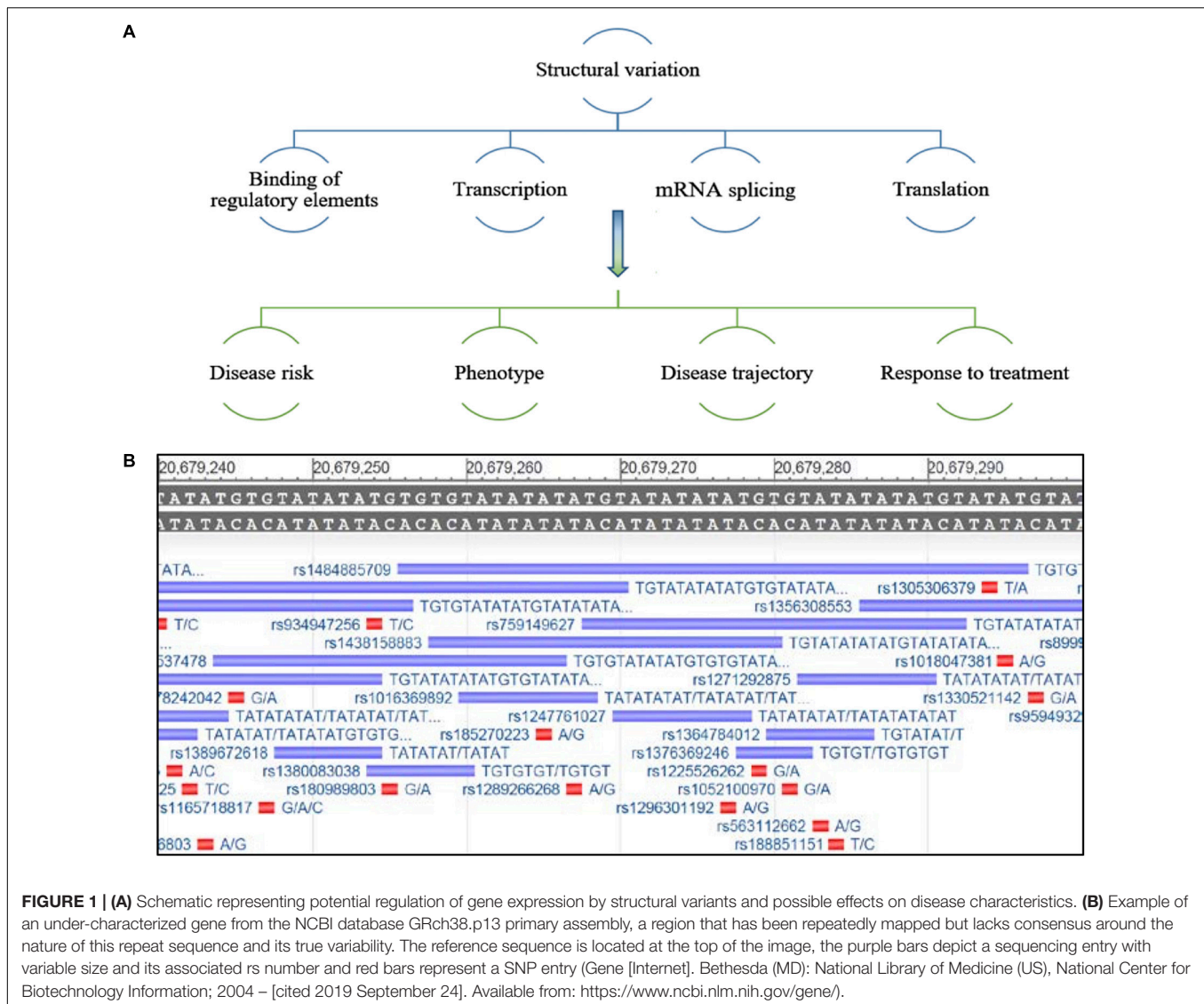
Historically, research on SVs has focused on genomic elements that are larger in size (>1 kb) and much easier to capture by high throughput techniques, such as copy number variations, transposable elements, larger insertion/deletions, translocations, and duplications (Sebat et al., 2004; Feuk et al., 2006; Huang et al., 2010; Alkan et al., 2011; Sudmant et al., 2015); yet other

small variable regions of the genome remain under characterized and are more difficult to capture with short-read sequencing techniques (**Figure 1B**; Chaisson et al., 2019; Ebbert et al., 2019). Recently, a comprehensive review highlighted the strengths and weakness of various sequencing methods and bioinformatics tools used for SV calling, and the difficulties in capturing smaller novel variations (Mahmoud et al., 2019). In addition, comparing the accuracy of SV characterization between methodologies when different *in silico* data sets are often utilized remains a challenge (Mahmoud et al., 2019). Furthermore, how these methods translate and generalize to patient samples is unclear, making it critical to establish standard procedures and benchmarks for the interpretation of SV data. There is a growing need to utilize SV information in the clinical setting to inform variation in patient phenotypes. Therefore we and others believe, important genes that may be critical to understanding the variation between patients that cannot be fully explained by SNPs or other mutations, should be reinvestigated for short SVs. Unrecognized or under characterized SVs could influence the expression of these genes, thereby contributing to the risk of ALS.

Recent studies investigating the transcriptome of the spinal cord anterior horn have identified significantly different RNA profiles between ALS patients and controls, for a multitude of gene pathways (D'Erchia et al., 2017). Since SVs can exhibit a range of regulatory effects that can impact levels of gene expression and potentially the phenotype, it is essential that these regions are also properly characterized. For example, a microsatellite repeat element in the promotor region of *PRPF31* (precursor mRNA-processing factor 31) results in some mutation carriers developing retinitis pigmentosa, whilst others remain asymptomatic (Rose et al., 2016). The length of this SV was shown to impact the penetrance of the mutation by suppressing transcription of this region by 50–115-fold, resulting in haploinsufficiency (Rose et al., 2016). Such disease mechanisms also warrant investigation in ALS.

STRUCTURAL VARIANTS IN ALS: *C9orf72* AND *ATXN2* VARIATION

An example of an SV that results in ALS pathogenesis is the repeat expansion in the *C9orf72* gene. The protein encoded by this gene is thought to play a role in endosomal membrane trafficking and autophagy (Farg et al., 2014). The SV region of *C9orf72* is a hexanucleotide repeat located in intron 1, GGGGCC that is usually repeated up to 30 times in healthy individuals. Expansion of this repeat to hundreds or thousands of repeated segments is a recognized cause of fALS, frontotemporal dementia, and occasionally also sALS (Mis et al., 2017). The DNA encoding this repeat is transcribed bi-directionally, resulting in nuclear RNA inclusions, and is thought to promote gain of function toxicity (Ly and Miller, 2018; Staats et al., 2019). Other potential mechanisms include *C9orf72* loss of function (Shi et al., 2018; Staats et al., 2019), or proteotoxicity (Gitler and Tsuji, 2016). In particular, patients with expanded repeats have a more severe phenotype, are predominantly bulbar onset, exhibit an earlier age of disease onset, cognitive and behavioral impairment and reduced survival (Byrne et al., 2012;

FIGURE 1 | (A) Schematic representing potential regulation of gene expression by structural variants and possible effects on disease characteristics. **(B)** Example of an under-characterized gene from the NCBI database GRch38.p13 primary assembly, a region that has been repeatedly mapped but lacks consensus around the nature of this repeat sequence and its true variability. The reference sequence is located at the top of the image, the purple bars depict a sequencing entry with variable size and its associated rs number and red bars represent a SNP entry (Gene [Internet]. Bethesda (MD): National Library of Medicine (US), National Center for Biotechnology Information; 2004 – [cited 2019 September 24]. Available from: https://www.ncbi.nlm.nih.gov/gene/).

Cooper-Knock et al., 2014). In addition, microsatellite repeats are subjected to an unconventional mechanism called repeat associated non-ATG (RAN) translation (Zu et al., 2011), whereby proteins can be coded by the additional two reading frames (GCA and AGC) resulting in up to six dipeptide proteins (Cleary and Ranum, 2014, 2017). The accumulation of these dipeptide proteins is seen in the central nervous system of individuals with ALS and contributes to a multitude of mechanisms that can induce protein-mediated toxicity (Balendra and Isaacs, 2018). One of the current strategies employed in the development of therapeutics for *C9orf72* is to target the SV region with antisense oligonucleotides to induce transcript degradation by RNase H enzymatic cleavage, preventing the build-up of toxic *C9orf72* transcript and protein. Other strategies focus on modulating the expression of transcription factors specifically involved in transcribing expanded repeats (Ly and Miller, 2018). Recently, it was shown that small ribosomal subunit protein (RPS25) plays a direct role in RAN translation, and decreasing its

levels through RNA interface mediated reduction prolonged the lifespan of *Drosophila* with the expanded repeat (Yamada et al., 2019). In addition, antisense oligonucleotide targeting of RPS25 enhanced the survival of *C9orf72* derived motor neurons reducing poly-GR and poly-PR foci (Yamada et al., 2019). The recent identification of a RAN translation regulator is a big step forward in demonstrating the functional implications of SV repeats in disease, and indicates antisense oligonucleotide or small molecules could be a viable therapeutic option for targeting RPS25 for patients with *C9orf72* (Hutten and Dormann, 2019). However, further research is needed to understand the regulation of expanded repeats and determine the relative contribution of repeat RNA and dipeptide repeat proteins to patient phenotype and cellular toxicity that promotes neurodegeneration in ALS.

Another SV contributing to ALS pathogenesis occurs in the gene encoding ataxin-2 (*ATXN2*). *ATXN2* contains a CAG repeat, initially found to be associated with a class of neuromuscular and neurological disorders, known as polyglutamine disorders,

caused by the expansion of the microsatellite repeat within the coding sequence (Al-Chalabi et al., 2012; Lattante et al., 2014). The ATXN2 protein is involved in endocytosis and modulates mTOR signals, critical to maintaining cell growth and survival, thereby modifying translation and mitochondrial function (Carmo-Silva et al., 2017). The N-terminal of this protein contains a polyglutamine tract that generally consists of 14–31 residues that when expanded, can carry up to 200 residues in the pathogenic state (Sproviero et al., 2017). Long expanded repeats were initially found to cause spinocerebellar ataxia 2 and subsequently, intermediate length repeats were shown to increase the risk of developing ALS (Daoud et al., 2011; Van Damme et al., 2011), with one study demonstrating that ALS risk increases exponentially with repeat length until the cut-off for developing spinocerebellar ataxia is reached (Sproviero et al., 2017). The polyglutamine disease causing mechanism differs between conditions and may include a loss of function resulting from hyper-methylation (Jin and Warren, 2000), a toxic gain of function through RAN translation (Scoles et al., 2015), protein misfolding and aggregation (Kayatekin et al., 2014), and in ALS, increasing TDP-43 toxicity (Elden et al., 2010). Longer repeats were expected to result in increased toxic TDP-43 build-up, resulting in increased risk of disease, an earlier age of onset and faster progression, however, this is not necessarily seen in patients (Al-Chalabi and Hardiman, 2013). The possibility of oligogenic inheritance is the likely explanation, where multiple risk factors from polymorphic structural variants are required to act together over time, with environmental stressors, to cause the development of ALS (Al-Chalabi et al., 2014), accounting for the particular variability seen in age-of-onset and disease progression. A better understanding of these variable regions of the genome and how they might work together to cumulatively increase disease risk, resulting in motor neuron dysfunction and susceptibility to neuronal degeneration, is imperative.

THERAPEUTIC DEVELOPMENT CHALLENGES

Over the past 20 years, more than 50 controlled trials of putative ALS therapeutics, testing 60 molecules have failed to demonstrate clinical efficacy (Petrov et al., 2017). Transgenic SOD1 mice have been used for the majority of ALS pre-clinical drug development studies; however, these models do not translate well to human disease (Mitsumoto et al., 2014). Alternative approaches need to be used to evaluate the potential efficacy of compounds. Identification and utilization of genetic markers, such as SVs informative for ALS, could be incorporated into clinical trial design to reduce the participant heterogeneity (Van Eijk et al., 2019).

Poor understanding of the pathogenic mechanisms of ALS neuro-degeneration remains a barrier to the development of novel therapeutic approaches. Moreover, there are few biomarkers that allow patient stratification according to disease mechanism (Agah et al., 2018; Khalil et al., 2018; Mitsumoto and Saito, 2018; Vejux et al., 2018). As a result, efficacy can only be evaluated by clinical measures during clinical trials

(Mitsumoto et al., 2014; Mitsumoto and Saito, 2018). Clearly, there is a critical unmet need to establish well-characterized molecular biomarkers that can be used as therapeutic targets, or to inform on the validity of certain treatment approaches. As ALS is a complex, heterogenous disorder with a varied age of onset and expression, it is likely that no single therapeutic will be effective for all patients. Therefore, we must develop strategies to identify patient subgroups and develop compounds to address the specific molecular defect.

Establishing molecular targets and markers for ALS could lead to improved patient stratification for clinical trials, to enable treatment effects to be identified within specific patient sub-groups. An example of the success of this approach is provided by clinical research with lithium carbonate in ALS patients. After a pilot study demonstrated attenuation of disease progression in a small number of ALS patients treated with lithium carbonate (Fornai et al., 2008), a number of follow up clinical studies have failed to replicate the finding (Aggarwal et al., 2010; Chiò et al., 2010; Miller et al., 2011; Verstraete et al., 2012; UKMND-LiCALS Study Group, 2013). In a meta-analysis of three trials that failed to show a significant effect of lithium treatment in ALS, Van Eijk et al. (2017) retrospectively demonstrated that lithium-treated patients who carried the UNC13A (C/C) genotype had a slower disease trajectory and showed a 70% improvement in 12 month survival, whilst carriers of the same genotype receiving no treatment had significantly reduced survival trajectories (Van Eijk et al., 2017). This survival benefit was only evident once the patients were stratified by their genotype, as the heterogeneous trial cohort originally masked the identification of any therapeutic benefit for a subgroup. In a more recent report Van Eijk et al. (2019) demonstrated that different genotypes including the repeat expansion C9orf72 can interact with both primary and secondary endpoints of clinical trials. Interestingly, in this report C9orf72 carriers did not have reduced survival, however, they did exhibit an accelerated monthly decline measured by ALSFRS compared to non-carriers. No pharmacogenetic interactions were demonstrated in the valproic acid trial, however, there was a pharmacognetic interaction between creatine treatment and the A allele of MOBP, whilst a dose response was observed for the C allele of UNC13A improving survival outcomes (Van Eijk et al., 2019). This highlights the importance of taking genetic information into account in clinical trials to enrich trial populations for potential responders. Identifying new genetic variations that may explain changes in gene expression in sALS patients will therefore be extremely useful to help inform both primary and secondary end points in clinical trial, and may improve the likelihood of clinical trial success.

CONCLUDING REMARKS

The methodologies currently used to elucidate ALS pathogenesis and inform drug development have not delivered effective therapeutic strategies to date. In our view, continuing to perform further genome wide association studies is unlikely to provide the breakthroughs that are urgently needed. Genome wide

studies can sometimes identify biochemical pathways involved in disease and indicate genes associated with fALS and sALS, however, in-depth characterization of these implicated regions may identify SVs that influence ALS susceptibility. Limitations of GWAS and even whole genome sequencing in identifying ALS risk must be recognized, since allele frequencies of variants or SNPs may not differ between patient cohorts and controls in these complex diseases. With increasing likelihood that SVs do indeed contribute to ALS risk, future investigations will need to incorporate SVs into genetic studies. It is possible that a combination of variants occurring frequently throughout healthy populations will collectively contribute to the vulnerability of motor neurons, and that this could be further exacerbated by both physiological and environmental insults. Particular SVs may better account for the variability in phenotypes and progression across ALS patient cohorts, and could be critical targets that can inform drug development. As our current molecular understanding of ALS has proven largely ineffective in easing the burden of ALS, clinical trials are likely to continue to fail if analyses are restricted to conventional strategies and platforms.

In silico investigations by our laboratory group reveal that there are a number of under-characterized genomic regions in ALS genes. Genetic data-bases including National Center for Biotechnology Information[1], Ensembl genome browser 97[2] and University of California Santa Cruz genome browser[3] have multiple sequence entries logged for genomic loci under different RS numbers. Sequencing technologies, whilst sometimes precise, have limited accuracy (Roses, 2016), as they only provide the location but not the specifics of each variant in individual patients. These poorly characterized regions of the genome could therefore contribute to the missing heritability of ALS. In most cases, the variation in allele length and allele frequency remains unclear, ultimately leading to the question, "how significant a role do SVs play in complex diseases, such as ALS?" We have focused on the development of accurate assays to genotype SVs. For example, investigation of the *SOD1* region has led to the characterization of SV1, a variant located within the 3′ untranslated region of *SCAF4*, a downstream gene that was previously overlooked. The function

[1] https://www.ncbi.nlm.nih.gov/gene/

[2] http://asia.ensembl.org/index.html

[3] https://genome.ucsc.edu/cgi-bin/hgGateway

of *SCAF4* has recently been elucidated, with the protein being necessary for accurate termination of transcription by ensuring correct polyadenylation site selection (Gregersen et al., 2019). This SV appears to influence *SOD1* expression, possibly through a toxic gain of function, and could more broadly influence ALS pathogenesis (Roses Allen, 2019). Continued investigations into the function of SV1 are presently underway in our laboratory. In addition, other variants appear to stratify sALS patients on the basis of survival and may in fact act as modifiers of gene expression (unpublished data). If this data is confirmed, it will not only indicate novel mechanisms contributing to ALS, but also allow patient stratification for enrichment of ALS clinical trials. Although SVs may not always be a viable drug target, they may indicate pathways that can be targeted to inform drug development. It is our belief that we need to re-assess these regions of the genome in order to identify some of the missing heritability of ALS and explain the phenotypic variability seen across this disease.

REFERENCES

Agah, E., Saleh, F., Moghaddam, H. S., Saghazadeh, A., Tafakhori, A., and Rezaei, N. (2018). CSF and blood biomarkers in amyotrophic lateral sclerosis: protocol for a systematic review and meta-analysis. *Syst. Rev.* 7:237. doi: 10.1186/s13643-018-0913-4

Aggarwal, S. P., Zinman, L., Simpson, E., McKinley, J., Jackson, K. E., Pinto, H., et al. (2010). Safety and efficacy of lithium in combination with riluzole for treatment of amyotrophic lateral sclerosis: a randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* 9, 481–488. doi: 10.1016/s1474-4422(10)70068-5

Al-Chalabi, A., Calvo, A., Chio, A., Colville, S., Ellis, C. M., Hardiman, O., et al. (2014). Analysis of amyotrophic lateral sclerosis as a multistep process: a

population-based modelling study. *Lancet Neurol.* 13, 1108–1113. doi: 10.1016/S1474-4422(14)70219-4

Al-Chalabi, A., Fang, F., Hanby, M. F., Leigh, P. N., Shaw, C. E., Ye, W., et al. (2010). An estimate of amyotrophic lateral sclerosis heritability using twin data. *J. Neurol. Neurosurg. Psychiatry* 81, 1324–1326. doi: 10.1136/jnnp.2010.207464

Al-Chalabi, A., and Hardiman, O. (2013). The epidemiology of ALS: a conspiracy of genes, environment and time. *Nat. Rev. Neurol.* 9, 617–628. doi: 10.1038/nrneurol.2013.203

Al-Chalabi, A., Jones, A., Troakes, C., King, A., Al-Sarraj, S., and van den Berg, L. H. (2012). The genetics and neuropathology of amyotrophic lateral sclerosis. *Acta Neuropathol.* 124, 339–352. doi: 10.1007/s00401-012-1022-4

Al-Chalabi, A., Van Den Berg, L. H., and Veldink, J. (2017). Gene discovery in amyotrophic lateral sclerosis: implications for clinical management. *Nat. Rev. Neurol.* 13, 96–104. doi: 10.1038/nrneurol.2016.182

Al-Chalabi, A., and Visscher, P. M. (2014). Motor neuron disease: common genetic variants and the heritability of ALS. *Nat. Rev. Neurol.* 10, 549–550. doi: 10.1038/nrneurol.2014.166

Alkan, C., Coe, B. P., and Eichler, E. E. (2011). Genome structural variation discovery and genotyping. *Nat. Rev. Genet.* 12, 363–376. doi: 10.1038/nrg2958

Al-Saif, A., Al-Mohanna, F., and Bohlega, S. (2011). A mutation in sigma-1 receptor causes juvenile amyotrophic lateral sclerosis. *Ann. Neurol.* 70, 913–919. doi: 10.1002/ana.22534

Andersen, P. M., and Al-Chalabi, A. (2011). Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nat. Rev. Neurol.* 7, 603–615. doi: 10.1038/nrneurol.2011.150

Arthur, K. C., Calvo, A., Price, T. R., Geiger, J. T., Chio, A., and Traynor, B. J. (2016). Projected increase in amyotrophic lateral sclerosis from 2015 to 2040. *Nat. Commun.* 7:12408. doi: 10.1038/ncomms12408

Balendra, R., and Isaacs, A. M. (2018). C9orf72-mediated ALS and FTD: multiple pathways to disease. *Nat. Rev. Neurol.* 14, 544–558. doi: 10.1038/s41582-018-0047-2

Bannwarth, S., Ait-El-Mkadem, S., Chaussenot, A., Genin, E. C., Lacas-Gervais, S., Fragaki, K., et al. (2014). A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain* 137, 2329–2345. doi: 10.1093/brain/awu138

Beck, J., Poulter, M., Hensman, D., Rohrer, J. D., Mahoney, C. J., Adamson, G., et al. (2013). Large C9orf72 hexanucleotide repeat expansions are seen in multiple neurodegenerative syndromes and are more frequent than expected in the UK population. *Am. J. Hum. Genet.* 92, 345–353. doi: 10.1016/j.ajhg.2013.01.011

Byrne, S., Elamin, M., Bede, P., Shatunov, A., Walsh, C., Corr, B., et al. (2012). 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. doi: 10.1016/S1474-4422(12)70014-5

Cameron, D. L., Di Stefano, L., and Papenfuss, A. T. (2019). Comprehensive evaluation and characterisation of short read general-purpose structural variant calling software. *Nat. Commun.* 10:3240. doi: 10.1038/s41467-019-11146-4

Carmo-Silva, S., Nobrega, C., de Almeida, L. P., and Cavadas, C. (2017). Unraveling the role of ataxin-2 in metabolism. *Trends Endocrinol. Metab.* 28, 309–318. doi: 10.1016/j.tem.2016.12.006

Chaisson, M. J., Sanders, A. D., Zhao, X., Malhotra, A., Porubsky, D., Rausch, T., et al. (2019). Multi-platform discovery of haplotype-resolved structural variation in human genomes. *Nat. Commun.* 10:1784. doi: 10.1038/s41467-018-08148-z

Chen, Y.-Z., Bennett, C. L., Huynh, H. M., Blair, I. P., Puls, I., Irobi, J., et al. (2004). DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am. J. Hum. Genet.* 74, 1128–1135. doi: 10.1086/421054

Chiang, C., Scott, A. J., Davis, J. R., Tsang, E. K., Li, X., Kim, Y., et al. (2017). The impact of structural variation on human gene expression. *Nat. Genet.* 49, 692–699. doi: 10.1038/ng.3834

Chiò, A., Borghero, G., Calvo, A., Capasso, M., Caponnetto, C., Corbo, M., et al. (2010). Lithium carbonate in amyotrophic lateral sclerosis: lack of efficacy in a dose-finding trial. *Neurology* 75, 619–625. doi: 10.1212/WNL.0b013e3181ed9e7c

Chow, C. Y., Landers, J. E., Bergren, S. K., Sapp, P. C., Grant, A. E., Jones, J. M., et al. (2009). Deleterious variants of FIG4, a phosphoinositide phosphatase, in patients with ALS. *Am. J. Hum. Genet.* 84, 85–88. doi: 10.1016/j.ajhg.2008.12.010

Cirulli, E. T., Lasseigne, B. N., Petrovski, S., Sapp, P. C., Dion, P. A., Leblond, C. S., et al. (2015). Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science* 347, 1436–1441. doi: 10.1126/science.aaa3650

Cleary, J. D., and Ranum, L. P. (2014). Repeat associated non-ATG (RAN) translation: new starts in microsatellite expansion disorders. *Curr. Opin. Genet. Dev.* 26, 6–15. doi: 10.1016/j.gde.2014.03.002

Cleary, J. D., and Ranum, L. P. (2017). New developments in RAN translation: insights from multiple diseases. *Curr. Opin. Genet. Dev.* 44, 125–134. doi: 10.1016/j.gde.2017.03.006

Cook, C., and Petrucelli, L. (2019). Genetic convergence brings clarity to the enigmatic red line in ALS. *Neuron* 101, 1057–1069. doi: 10.1016/j.neuron.2019.02.032

Cooper-Knock, J., Shaw, P. J., and Kirby, J. (2014). The widening spectrum of C9ORF72-related disease; genotype/phenotype correlations and potential modifiers of clinical phenotype. *Acta Neuropathol.* 127, 333–345. doi: 10.1007/s00401-014-1251-9

Corcia, P., Mayeux-Portas, V., Khoris, J., de Toffol, B., Autret, A., Müh, J. P., et al. (2002). Abnormal SMN1 gene copy number is a susceptibility factor for amyotrophic lateral sclerosis. *Ann. Neurol.* 51, 243–246. doi: 10.1002/ana.10104

Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G., et al. (1993). Gene dose of apolipoprotein E type 4 allele and the risk of Alzheimer's disease in late onset families. *Science* 261, 921–923. doi: 10.1126/science.8346443

Couthouis, J., Hart, M. P., Erion, R., King, O. D., Diaz, Z., Nakaya, T., et al. (2012). Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. *Hum. Mol. Genet.* 21, 2899–2911. doi: 10.1093/hmg/dds116

Daoud, H., Belzil, V., Martins, S., Sabbagh, M., Provencher, P., Lacomblez, L., et al. (2011). Association of long ATXN2 CAG repeat sizes with increased risk of amyotrophic lateral sclerosis. *Arch. Neurol.* 68, 739–742. doi: 10.1001/archneurol.2011.111

DeJesus-Hernandez, M., Mackenzie, I. R., Boeve, B. F., Boxer, A. L., Baker, M., Rutherford, N. J., et al. (2011). Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* 72, 245–256. doi: 10.1016/j.neuron.2011.09.011

Deng, H.-X., Chen, W., Hong, S.-T., Boycott, K. M., Gorrie, G. H., Siddique, N., et al. (2011). Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* 477, 211–215. doi: 10.1038/nature10353

D'Erchia, A. M., Gallo, A., Manzari, C., Raho, S., Horner, D. S., Chiara, M., et al. (2017). Massive transcriptome sequencing of human spinal cord tissues provides new insights into motor neuron degeneration in ALS. *Sci. Rep.* 7:10046. doi: 10.1038/s41598-017-10488-7

Dharmadasa, T., Henderson, R. D., Talman, P. S., Macdonell, R. A., Mathers, S., Schultz, D. W., et al. (2017). Motor neurone disease: progress and challenges. *Med. J. Aust.* 206, 357–362.

Doble, A. (1996). The pharmacology and mechanism of action of riluzole. *Neurology* 47(6 Suppl. 4), 233S–241S. doi: 10.1212/wnl.47.6_suppl_4.233s

Droppelmann, C. A., Wang, J., Campos-Melo, D., Keller, B., Volkening, K., Hegele, R. A., et al. (2013). Detection of a novel frameshift mutation and regions with homozygosis within ARHGEF28 gene in familial amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Frontotemporal Degener.* 14, 444–451. doi: 10.3109/21678421.2012.758288

Ebbert, M. T. W., Jensen, T. D., Jansen-West, K., Sens, J. P., Reddy, J. S., Ridge, P. G., et al. (2019). Systematic analysis of dark and camouflaged genes reveals disease-relevant genes hiding in plain sight. *Genome Biol.* 20:97. doi: 10.1186/s13059-019-1707-2

Elden, A. C., Kim, H.-J., Hart, M. P., Chen-Plotkin, A. S., Johnson, B. S., Fang, X., et al. (2010). Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* 466, 1069–1075. doi: 10.1038/nature09320

Farg, M. A., Sundaramoorthy, V., Sultana, J. M., Yang, S., Atkinson, R. A., Levina, V., et al. (2014). C9ORF72, implicated in amyotrophic lateral sclerosis and frontotemporal dementia, regulates endosomal trafficking. *Hum. Mol. Genet.* 23, 3579–3595. doi: 10.1093/hmg/ddu068

Fecto, F., Yan, J., Vemula, S. P., Liu, E., Yang, Y., Chen, W., et al. (2011). SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *JAMA Neurol.* 68, 1440–1446.

Feuk, L., Carson, A. R., and Scherer, S. W. (2006). Structural variation in the human genome. *Nat. Rev. Genet.* 7, 85–97.

Figlewicz, D. A., Rouleau, G. A., Krizus, A., and Julien, J.-P. (1993). Polymorphism in the multi-phosphorylation domain of the human neurofilament heavy-subunit-encoding gene. *Gene* 132, 297–300. doi: 10.1016/0378-1119(93)90211-k

Fornai, F., Longone, P., Cafaro, L., Kastsiuchenka, O., Ferrucci, M., Manca, M. L., et al. (2008). Lithium delays progression of amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. U.S.A.* 105, 2052–2057. doi: 10.1073/pnas.0708022105

Freischmidt, A., Wieland, T., Richter, B., Ruf, W., Schaeffer, V., Müller, K., et al. (2015). Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nat. Neurosci.* 18, 631–636. doi: 10.1038/nn.4000

Gitcho, M. A., Baloh, R. H., Chakraverty, S., Mayo, K., Norton, J. B., Levitch, D., et al. (2008). TDP-43 A315T mutation in familial motor neuron disease. *Ann. Neurol.* 63, 535–538.

Gitler, A. D., and Tsuiji, H. (2016). There has been an awakening: emerging mechanisms of C9orf72 mutations in FTD/ALS. *Brain Res.* 1647, 19–29. doi: 10.1016/j.brainres.2016.04.004

Greenway, M. J., Alexander, M. D., Ennis, S., Traynor, B. J., Corr, B., Frost, E., et al. (2004). A novel candidate region for ALS on chromosome 14q11.2. *Neurology* 63, 1936–1938. doi: 10.1212/01.wnl.0000144344.39103.f6

Greenway, M. J., Andersen, P. M., Russ, C., Ennis, S., Cashman, S., Donaghy, C., et al. (2006). ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. *Nat. Genet.* 38, 411–413. doi: 10.1038/ng1742

Gregersen, L. H., Mitter, R., Ugalde, A. P., Nojima, T., Proudfoot, N. J., Agami, R., et al. (2019). SCAF4 and SCAF8, mRNA anti-terminator proteins. *Cell* 177, 1797–1813.e18. doi: 10.1016/j.cell.2019.04.038

Hadano, S., Hand, C. K., Osuga, H., Yanagisawa, Y., Otomo, A., Devon, R. S., et al. (2001). A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nat. Genet.* 29, 166–173. doi: 10.1038/ng1001-166

Huang, C. R. L., Schneider, A. M., Lu, Y., Niranjan, T., Shen, P., Robinson, M. A., et al. (2010). Mobile interspersed repeats are major structural variants in the human genome. *Cell* 141, 1171–1182. doi: 10.1016/j.cell.2010.05.026

Hutten, S., and Dormann, D. (2019). RAN translation down. *Nat. Neurosci.* 22, 1379–1380. doi: 10.1038/s41593-019-0482-4

Jin, P., and Warren, S. T. (2000). Understanding the molecular basis of fragile X syndrome. *Hum. Mol. Genet.* 9, 901–908. doi: 10.1093/hmg/9.6.901

Johnson, J. O., Mandrioli, J., Benatar, M., Abramzon, Y., Van Deerlin, V. M., Trojanowski, J. Q., et al. (2010). Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* 68, 857–864. doi: 10.1016/j.neuron.2010.11.036

Johnson, J. O., Pioro, E. P., Boehringer, A., Chia, R., Feit, H., Renton, A. E., et al. (2014). Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat. Neurosci.* 17, 664–666. doi: 10.1038/nn.3688

Kayatekin, C., Matlack, K. E., Hesse, W. R., Guan, Y., Chakrabortee, S., Russ, J., et al. (2014). Prion-like proteins sequester and suppress the toxicity of huntingtin exon 1. *Proc. Natl. Acad. Sci. U.S.A.* 111, 12085–12090. doi: 10.1073/pnas.1412504111

Kenna, K. P., van Doormaal, P. T. C., Dekker, A. M., Ticozzi, N., Kenna, B. J., Diekstra, F. P., et al. (2016). NEK1 variants confer susceptibility to amyotrophic lateral sclerosis. *Nat. Genet.* 48, 1037–1042. doi: 10.1038/ng.3626

Khalil, M., Teunissen, C. E., Otto, M., Piehl, F., Sormani, M. P., Gattringer, T., et al. (2018). Neurofilaments as biomarkers in neurological disorders. *Nat. Rev. Neurol.* 14, 577–589. doi: 10.1038/s41582-018-0058-z

Kim, H. J., Kim, N. C., Wang, Y.-D., Scarborough, E. A., Moore, J., Diaz, Z., et al. (2013). Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495, 467–473. doi: 10.1038/nature11922

Kinsley, L., and Siddique, T. (2015). "Amyotrophic lateral sclerosis overview," in *GeneReviews® [Internet]*, eds R. A. Pagon, M. P. Adam, H. H. Ardinger, S. E. Wallace, A. Amemiya, L. J. H. Bean, et al. (Seattle, WA: University of Washington).

Lambrechts, D., Storkebaum, E., Morimoto, M., Del-Favero, J., Desmet, F., Marklund, S. L., et al. (2003). VEGF is a modifier of amyotrophic lateral sclerosis in mice and humans and protects motoneurons against ischemic death. *Nat. Genet.* 34, 383–394. doi: 10.1038/ng1211

Landers, J. E., Melki, J., Meininger, V., Glass, J. D., Van Den Berg, L. H., Van Es, M. A., et al. (2009). Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci.* 106, 9004–9009. doi: 10.1073/pnas.0812937106

Lattante, S., Millecamps, S., Stevanin, G., Rivaud-Péchoux, S., Moigneu, C., Camuzat, A., et al. (2014). Contribution of ATXN2 intermediary polyQ expansions in a spectrum of neurodegenerative disorders. *Neurology* 83, 990–995. doi: 10.1212/WNL.0000000000000778

Leung, C. L., He, C. Z., Kaufmann, P., Chin, S. S., Naini, A., Liem, R. K. H., et al. (2004). A pathogenic peripherin gene mutation in a patient with amyotrophic

lateral sclerosis. *Brain Pathol.* 14, 290–296. doi: 10.1111/j.1750-3639.2004.tb00066.x

Ly, C. V., and Miller, T. M. (2018). Emerging antisense oligonucleotide and viral therapies for amyotrophic lateral sclerosis. *Curr. Opin. Neurol.* 31, 648–654. doi: 10.1097/WCO.0000000000000594

Lyall, D. M., Royle, N. A., Harris, S. E., Bastin, M. E., Maniega, S. M., Murray, C., et al. (2013). Alzheimer's disease susceptibility genes APOE and TOMM40, and hippocampal volumes in the Lothian birth cohort 1936. *PLoS One* 8:e80513. doi: 10.1371/journal.pone.0080513

Mahmoud, M., Gobet, N., Cruz-Dávalos, D. I., Mounier, N., Dessimoz, C., and Sedlazeck, F. J. (2019). Structural variant calling: the long and the short of it. *Genome Biol.* 20:246. doi: 10.1186/s13059-019-1828-7

Maruyama, H., Morino, H., Ito, H., Izumi, Y., Kato, H., Watanabe, Y., et al. (2010). Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* 465, 223–226. doi: 10.1038/nature08971

Mejzini, R., Flynn, L. L., Pitout, I. L., Fletcher, S., Wilton, S. D., and Akkari, P. A. (2019). ALS genetics, mechanisms, and therapeutics: where are we now? *Front. Neurosci.* 13:1310. doi: 10.3389/fnins.2019.01310

Miller, R., Moore, D., Forshew, D., Katz, J., Barohn, R., Valan, M., et al. (2011). Phase II screening trial of lithium carbonate in amyotrophic lateral sclerosis: examining a more efficient trial design. *Neurology* 77, 973–979. doi: 10.1212/WNL.0b013e31822dc7a5

Mis, M. S. C., Brajkovic, S., Tafuri, F., Bresolin, N., Comi, G. P., and Corti, S. (2017). Development of therapeutics for C9ORF72 ALS/FTD-related disorders. *Mol. Neurobiol.* 54, 4466–4476. doi: 10.1007/s12035-016-9993-0

Mitchell, J., Paul, P., Chen, H.-J., Morris, A., Payling, M., Falchi, M., et al. (2010). Familial amyotrophic lateral sclerosis is associated with a mutation in D-amino acid oxidase. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7556–7561.

Mitsumoto, H., Brooks, B. R., and Silani, V. (2014). Clinical trials in amyotrophic lateral sclerosis: why so many negative trials and how can trials be improved? *Lancet Neurol.* 13, 1127–1138. doi: 10.1016/S1474-4422(14)70129-2

Mitsumoto, H., and Saito, T. (2018). A prognostic biomarker in amyotrophic lateral sclerosis. *Clin. Neurol.* 58, 729–736. doi: 10.5692/clinicalneurol.cn-001220

Morita, M., Al-Chalabi, A., Andersen, P., Hosler, B., Sapp, P., Englund, E., et al. (2006). A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology* 66, 839–844. doi: 10.1212/01.wnl.0000200048.53766.b4

Münch, C., Sedlmeier, R., Meyer, T., Homberg, V., Sperfeld, A., Kurt, A., et al. (2004). Point mutations of the p150 subunit of dynactin (DCTN1) gene in ALS. *Neurology* 63, 724–726. doi: 10.1212/01.wnl.0000134608.83927.b1

Nguyen, H. P., Van Broeckhoven, C., and van der Zee, J. (2018). ALS genes in the genomic era and their implications for FTD. *Trends Genet.* 34, 404–423. doi: 10.1016/j.tig.2018.03.001

Nicolas, A., Kenna, K. P., Renton, A. E., Ticozzi, N., Faghri, F., Chia, R., et al. (2018). Genome-wide analyses identify KIF5A as a novel ALS gene. *Neuron* 97, 1268–1283.e6. doi: 10.1016/j.neuron.2018.02.027

Nishimura, A. L., Mitne-Neto, M., Silva, H. C., Richieri-Costa, A. N., Middleton, S., Cascio, D., et al. (2004). A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. *Am. J. Hum. Genet.* 75, 822–831. doi: 10.1086/425287

Okamoto, K., Hirai, S., Amari, M., Watanabe, M., and Sakurai, A. (1993). Bunina bodies in amyotrophic lateral sclerosis immunostained with rabbit anti-cystatin C serum. *Neurosci. Lett.* 162, 125–128. doi: 10.1016/0304-3940(93)90576-7

Orlacchio, A., Babalini, C., Borreca, A., Patrono, C., Massa, R., Basaran, S., et al. (2010). SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis. *Brain* 133, 591–598. doi: 10.1093/brain/awp325

Parkinson, N., Ince, P., Smith, M., Highley, R., Skibinski, G., Andersen, P., et al. (2006). ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). *Neurology* 67, 1074–1077. doi: 10.1212/01.wnl.0000231510.89311.8b

Petrov, D., Mansfield, C., Moussy, A., and Hermine, O. (2017). ALS clinical trials review: 20 years of failure. Are we any closer to registering a new treatment? *Front. Aging Neurosci.* 9:68. doi: 10.3389/fnagi.2017.00068

Renton, A. E., Chiò, A., and Traynor, B. J. (2014). State of play in amyotrophic lateral sclerosis genetics. *Nat. Neurosci.* 17, 17–23. doi: 10.1038/nn.3584

Renton, A. E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J. R., et al. (2011). A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* 72, 257–268. doi: 10.1016/j.neuron.2011.09.010

Robberecht, W., and Philips, T. (2013). The changing scene of amyotrophic lateral sclerosis. *Nat. Rev. Neurosci.* 14, 248–264. doi: 10.1038/nrn3430

Rose, A. M., Shah, A. Z., Venturini, G., Krishna, A., Chakravarti, A., Rivolta, C., et al. (2016). Transcriptional regulation of PRPF31 gene expression by MSR1 repeat elements causes incomplete penetrance in retinitis pigmentosa. *Sci. Rep.* 6:19450. doi: 10.1038/srep19450

Rosen, D. R., Siddique, T., Patterson, D., Figlewicz, D. A., Sapp, P., Hentati, A., et al. (1993). Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* 362, 59–62.

Roses, A. (2016). Polyallelic structural variants can provide accurate, highly informative genetic markers focused on diagnosis and therapeutic targets: accuracy vs. Precision. *Clin. Pharmacol. Ther.* 99, 169–171. doi: 10.1002/cpt.288

Roses, A., Lutz, M., Amrine-Madsen, H., Saunders, A., Crenshaw, D., Sundseth, S., et al. (2010). A TOMM40 variable-length polymorphism predicts the age of late-onset Alzheimer's disease. *Pharmacogenomics J.* 10, 375–384. doi: 10.1038/tpj.2009.69

Roses, A. D., Akkari, P. A., Chiba-Falek, O., Lutz, M. W., Gottschalk, W. K., Saunders, A. M., et al. (2016). Structural variants can be more informative for disease diagnostics, prognostics and translation than current SNP mapping and exon sequencing. *Expert Opin. Drug Metab. Toxicol.* 12, 135–147. doi: 10.1517/17425255.2016.1133586

Roses Allen, D. (2019). Methods for detecting structural variants in neurodegenerative disease. US Patent No 201,716,308,315.

Rothstein, J. D. (2017). Edaravone: a new drug approved for ALS. *Cell* 171:725. doi: 10.1016/j.cell.2017.10.011

Saul, R., Lutz, M. W., Burns, D. K., Roses, A. D., and Chiba-Falek, O. (2016). The SSV evaluation system: a tool to prioritize short structural variants for studies of possible regulatory and causal variants. *Hum. Mutat.* 37, 877–883. doi: 10.1002/humu.23023

Schymick, J. C., Yang, Y., Andersen, P. M., Vonsattel, J. P., Greenway, M., Momeni, P., et al. (2007). Progranulin mutations and amyotrophic lateral sclerosis or amyotrophic lateral sclerosis-frontotemporal dementia phenotypes. *J. Neurol. Neurosurg. Psychiatry* 78, 754–756.

Scoles, D. R., Ho, M. H., Dansithong, W., Pflieger, L. T., Petersen, L. W., Thai, K. K., et al. (2015). Repeat associated non-AUG translation (RAN translation) dependent on sequence downstream of the ATXN2 CAG repeat. *PLoS One* 10:e0128769. doi: 10.1371/journal.pone.0128769

Sebat, J., Lakshmi, B., Troge, J., Alexander, J., Young, J., Lundin, P., et al. (2004). Large-scale copy number polymorphism in the human genome. *Science* 305, 525–528. doi: 10.1126/science.1098918

Shi, Y., Lin, S., Staats, K. A., Li, Y., Chang, W.-H., Hung, S.-T., et al. (2018). Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons. *Nat. Med.* 24, 313–325. doi: 10.1038/nm.4490

Simpson, C. L., Lemmens, R., Miskiewicz, K., Broom, W. J., Hansen, V. K., van Vught, P. W., et al. (2008). Variants of the elongator protein 3 (ELP3) gene are associated with motor neuron degeneration. *Hum. Mol. Genet.* 18, 472–481. doi: 10.1093/hmg/ddn375

Slowik, A., Tomik, B., Wolkow, P., Partyka, D., Turaj, W., Malecki, M., et al. (2006). Paraoxonase gene polymorphisms and sporadic ALS. *Neurology* 67, 766–770. doi: 10.1212/01.wnl.0000219565.32247.11

Smith, B. N., Ticozzi, N., Fallini, C., Gkazi, A. S., Topp, S., Kenna, K. P., et al. (2014). Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS. *Neuron* 84, 324–331. doi: 10.1016/j.neuron.2014.09.027

Smith, B. N., Topp, S. D., Fallini, C., Shibata, H., Chen, H.-J., Troakes, C., et al. (2017). Mutations in the vesicular trafficking protein annexin A11 are associated with amyotrophic lateral sclerosis. *Sci. Transl. Med.* 9:eaad9157. doi: 10.1126/scitranslmed.aad9157

Sproviero, W., Shatunov, A., Stahl, D., Shoai, M., van Rheenen, W., Jones, A. R., et al. (2017). ATXN2 trinucleotide repeat length correlates with risk of ALS. *Neurobiol. Aging* 51, 178.e1–178.e9. doi: 10.1016/j.neurobiolaging.2016.11.010

Staats, K., Seah, C., Sahimi, A., Wang, Y., Koutsodendris, N., Lin, S., et al. (2019). Small molecule inhibition of PIKFYVE kinase rescues gain-and loss-of-function C9ORF72 ALS/FTD disease processes in vivo. *bioRxiv [Preprint]* doi: 10.1101/685800

Staats, K. A., Van Helleputte, L., Jones, A. R., Bento-Abreu, A., Van Hoecke, A., Shatunov, A., et al. (2013). Genetic ablation of phospholipase C delta 1 increases survival in SOD1G93A mice. *Neurobiol. Dis.* 60, 11–17. doi: 10.1016/j.nbd.2013.08.006

Sudmant, P. H., Rausch, T., Gardner, E. J., Handsaker, R. E., Abyzov, A., Huddleston, J., et al. (2015). An integrated map of structural variation in 2,504 human genomes. *Nature* 526, 75–81. doi: 10.1038/nature15394

Talman, P., Duong, T., Vucic, S., Mathers, S., Venkatesh, S., Henderson, R., et al. (2016). Identification and outcomes of clinical phenotypes in amyotrophic lateral sclerosis/motor neuron disease: Australian national motor neuron disease observational cohort. *Br. Med. J. Open* 6:e012054. doi: 10.1136/bmjopen-2016-012054

Ticozzi, N., Vance, C., Leclerc, A., Keagle, P., Glass, J., McKenna-Yasek, D., et al. (2011). Mutational analysis reveals the FUS homolog TAF15 as a candidate gene for familial amyotrophic lateral sclerosis. *Am. J. Med. Genet. Part B Neuropsychiatr. Genet.* 156, 285–290. doi: 10.1002/ajmg.b.31158

UKMND-LiCALS Study Group (2013). Lithium in patients with amyotrophic lateral sclerosis (LiCALS): a phase 3 multicentre, randomised, double-blind, placebo-controlled trial. *Lancet Neurol.* 12, 339–345. doi: 10.1016/S1474-4422(13)70037-1

Van Damme, P., Veldink, J. H., van Blitterswijk, M., Corveleyn, A., van Vught, P. W., Thijs, V., et al. (2011). Expanded ATXN2 CAG repeat size in ALS identifies genetic overlap between ALS and SCA2. *Neurology* 76, 2066–2072. doi: 10.1212/WNL.0b013e31821f445b

Van Eijk, R. P., Eijkemans, M. J., Nikolakopoulos, S., Jansen, M. D., Westeneng, H.-J., Van Eijk, K. R., et al. (2019). Pharmacogenetic interactions in amyotrophic lateral sclerosis: a step closer to a cure? *Pharmacogenomics J.* doi: 10.1038/s41397-019-0111-3 [Epub ahead of print].

Van Eijk, R. P., Jones, A. R., Sproviero, W., Shatunov, A., Shaw, P. J., Leigh, P. N., et al. (2017). Meta-analysis of pharmacogenetic interactions in amyotrophic lateral sclerosis clinical trials. *Neurology* 89, 1915–1922. doi: 10.1212/WNL.0000000000004606

Van Es, M. A., Van Vught, P. W., Blauw, H. M., Franke, L., Saris, C. G., Andersen, P. M., et al. (2007). ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. *Lancet Neurol.* 6, 869–877.

Van Es, M. A., Van Vught, P. W., Blauw, H. M., Franke, L., Saris, C. G., Van Den Bosch, L., et al. (2008). Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis. *Nat. Genet.* 40, 29–31.

Van Es, M. A., Veldink, J. H., Saris, C. G., Blauw, H. M., van Vught, P. W., Birve, A., et al. (2009). Genome-wide association study identifies 19p13. 3 (UNC13A) and 9p21. 2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat. Genet.* 41, 1083–1087. doi: 10.1038/ng.442

Van Rheenen, W., Shatunov, A., Dekker, A. M., McLaughlin, R. L., Diekstra, F. P., Pulit, S. L., et al. (2016). Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nat. Genet.* 48, 1043–1048. doi: 10.1038/ng.3622

Vance, C., Rogelj, B., Hortobágyi, T., De Vos, K. J., Nishimura, A. L., Sreedharan, J., et al. (2009). Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* 323, 1208–1211. doi: 10.1126/science.1165942

Vejux, A., Namsi, A., Nury, T., Moreau, T., and Lizard, G. (2018). Biomarkers of amyotrophic lateral sclerosis: current status and interest of oxysterols and phytosterols. *Front. Mol. Neurosci.* 11:12. doi: 10.3389/fnmol.2018.00012

Verstraete, E., Veldink, J. H., Huisman, M. H., Draak, T., Uijtendaal, E. V., van der Kooi, A. J., et al. (2012). Lithium lacks effect on survival in amyotrophic lateral sclerosis: a phase IIb randomised sequential trial. *J. Neurol. Neurosurg. Psychiatry* 83, 557–564. doi: 10.1136/jnnp-2011-302021

Wang, X.-S., Lee, S., Simmons, Z., Boyer, P., Scott, K., Liu, W., et al. (2004). Increased incidence of the Hfe mutation in amyotrophic lateral sclerosis and related cellular consequences. *J. Neurol. Sci.* 227, 27–33. doi: 10.1016/j.jns.2004.08.003

Williams, K. L., Topp, S., Yang, S., Smith, B., Fifita, J. A., Warraich, S. T., et al. (2016). CCNF mutations in amyotrophic lateral sclerosis and frontotemporal dementia. *Nat. Commun.* 7:11253. doi: 10.1038/ncomms11253

Wu, C.-H., Fallini, C., Ticozzi, N., Keagle, P. J., Sapp, P. C., Piotrowska, K., et al. (2012). Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature* 488, 499–503. doi: 10.1038/nature11280

Yamada, S. B., Gendron, T. F., Niccoli, T., Genuth, N. R., Grosely, R., Shi, Y., et al. (2019). RPS25 is required for efficient RAN translation of C9orf72 and other neurodegenerative disease-associated nucleotide repeats. *Nat. Neurosci.* 22, 1383–1388. doi: 10.1038/s41593-019-0455-7

Zu, T., Gibbens, B., Doty, N. S., Gomes-Pereira, M., Huguet, A., Stone, M. D., et al. (2011). Non-ATG–initiated translation directed by microsatellite expansions. *Proc. Natl. Acad. Sci. U.S.A.* 108, 260–265. doi: 10.1073/pnas.1013343108