Amyotrophic Lateral Sclerosis Genetic Studies: From Genome-wide Association Mapping to Genome Sequencing

Ji He[1,2,3], Marie Mangelsdorf[1], Dongsheng Fan[2], Perry Bartlett[1], and Matthew A. Brown[3]

Abstract

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease of obscure etiology. Multiple genetic studies have been conducted to advance our understanding of the disease, employing a variety of techniques such as linkage mapping in families, to genome-wide association studies and sequencing based approaches such as whole exome sequencing and whole genome sequencing and a few epigenetic analyses. While major progress has been made, the majority of the genetic variation involved in ALS is yet to be undefined. The optimal study designs to investigate ALS depend on the genetic model for the disease, and it is likely that different approaches will be required to map genes involved in familial and sporadic disease. The potential approaches and their strengths and weaknesses are discussed.

Keywords

amyotrophic lateral sclerosis, genetic study, epigenetic study, genome-wide association study, next-generation sequencing study

Introduction

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease caused by selective loss of upper and lower motor neurons (common clinical manifestations of ALS are shown in Fig. 1). The disease typically has a late onset and is often fatal within 3 to 5 years. No effective treatments exist for ALS; the only current treatment is riluzole, which is thought to affect glutamate metabolism and only modestly extends survival time (Miller and others 2012). The worldwide incidence of ALS is approximately 0.3 to 7.0 new cases per 100,000 each year (Cronin and others 2007). Nearly 10% of ALS cases are classified as familial ALS (fALS) (Byrne and others 2011), wherein a family history of the disease is known. In these families, the pattern is primarily consistent with autosomal dominant inheritance, although other hereditary patterns have been found (Siddique and Ajroud-Driss 2011). Approximately half of familial cases can be explained by specific genes, the majority of which appear to be rare highly penetrant de novo mutations in each family involved. The 90% of ALS cases where no family history is present are termed sporadic ALS (sALS). The genetic etiology of both sALS and nearly ~30% of fALS (European descent) remains obscure. Whether sALS is primarily polygenic, because of highly penetrant de novo mutations or inherited mutations with low penetrance is unclear, and this knowledge will be critical in the design of future gene-mapping studies. Detailed information regarding ALS-related genes is available via the amyotrophic lateral sclerosis online genetics database (http://alsod.iop.kcl.ac.uk/), the ALS mutation database (https://reseq.lifesciencedb.jp/resequence/SearchDisease.do?targetId=1), and the ALSGene database (http://www.alsgene.org).

Exploring the genetic etiology of ALS has provided fundamental insights into the process underlying neuron degeneration and majority of our current knowledge of ALS. On this basis, considerable efforts have been devoted to unravel pathophysiological mechanisms.

[1]Queensland Brain Institute, University of Queensland, St. Lucia, Brisbane, Australia
[2]Department of Neurology, Peking University Third Hospital, Beijing, China
[3]University of Queensland Diamantina Institute, Translational Research Institute, Princess Alexandra Hospital, Woolloongabba, Brisbane, Australia

Corresponding Author:
Matthew A. Brown , Translational Research Institute Pty Ltd, 37 Kent Street, Woolloongabba, Brisbane, Queensland 4102, Australia.
Email: matt.brown@uq.edu.au

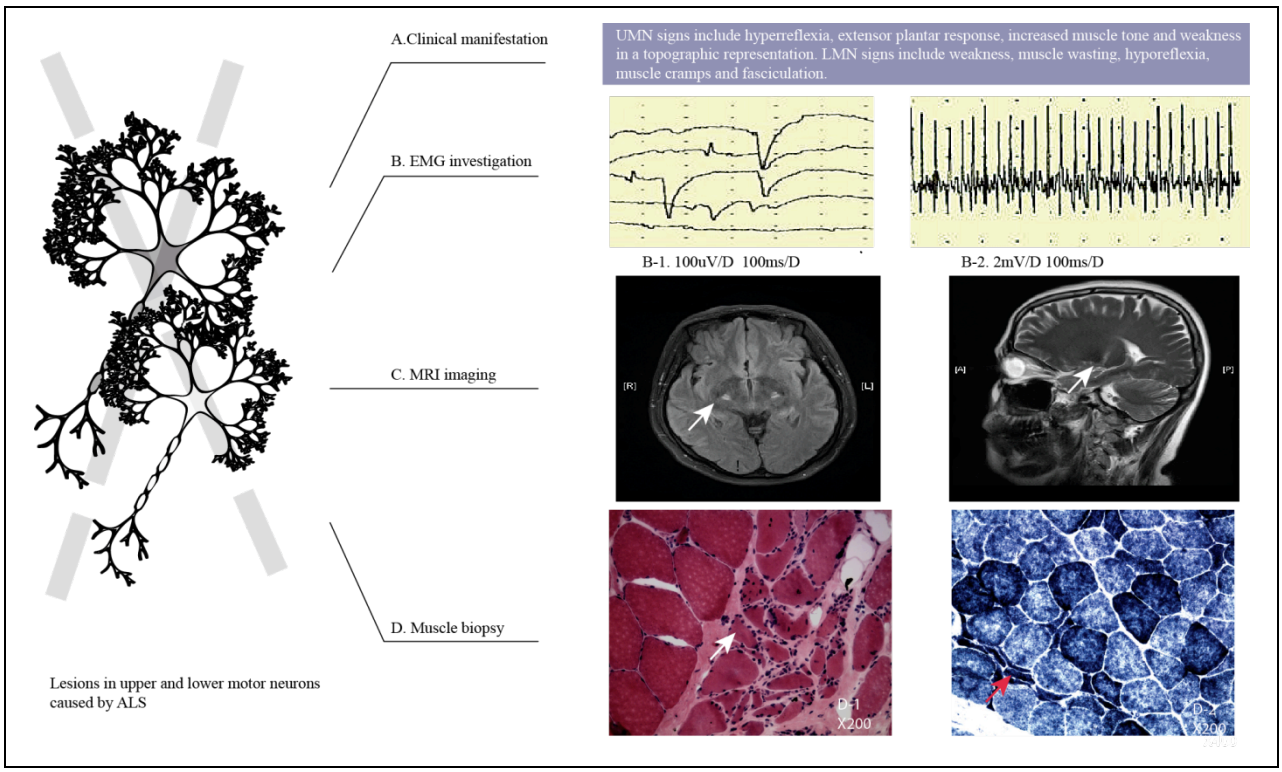

Figure 1. Common clinical methods and manifestations in amyotrophic lateral sclerosis (ALS). (A) LMN = lower motor neurons; UMN = upper motor neurons. (B) EMG = electromyography: Nerve conduction study and needle electromyography remain the most important diagnostic testing for ALS. The former is used primarily to help rule out other disorders and the latter to establish evidence for widespread active denervation and chronic reinnervation. Evidence for denervation in ALS on needle EMG includes (B-1)—fibrillation potentials and positive sharp waves and (B-2)—pathologic interference pattern. (C) This MRI demonstrates increased T2 signal within the posterior part of the internal capsule and can be tracked to the subcortical white matter of the motor cortex, outlining the corticospinal tract. (D) Routine muscle biopsy: ALS (D-1—hematoxylin and eosin frozen section 200× and D-2—NADH stain 200×) may show grouped atrophy; D-2—Small muscle fibers stain darkly on NADH.

Multiple pathogenic processes have been reported that support the view of multiple routes to a common endpoint of progressive upper motor neuron and lower motor neuron loss (Fig. 2).

Since the discovery of mutations in the *SOD1* gene causing fALS in 1993(Rosen and others 1993), a growing number of causative genes and related loci have been identified (Table 1), including *TARDBP*, *FUS* and *C9orf72*. A breakthrough in the understanding of ALS pathogenesis was the discovery that TDP-43 is a central component of the ubiquitin-positive neuronal inclusions that are the pathological hallmark in ALS (Neumann and others 2006). TDP-43 is an RNA-binding protein that is encoded by the gene *TARDBP* and mutations were subsequently found in rare ALS case. The discovery of ALS causing mutations in several other genes whose encoded proteins are involved in RNA processing has further implicated this molecular pathway in the pathogenesis in ALS. These include mutations in *FUS*, encoding an RNA-binding protein with similarity to TDP-43 (Kwiatkowski and others 2009; Vance and others 2009) and a repeat expansion in *C9orf72*, associated with RNA foci in cells from ALS cases (DeJesus-Hernandez and others 2011). A growing number of genes have been implicated in ALS by discovery of rare sequence variants in ALS cases, including *VCP* (Wu and others 2012) and *PFN1* (Johnson and others 2010), although molecular evidence supporting a role as an ALS causative gene is not entirely convincing for all genes and requires further validation.

In order to advance our understanding of the aetiopathogenesis of ALS, extensive candidate gene and hypothesis-free gene-mapping studies have been pursued. In recent years, revolutions in genetic techniques have taken place. The development of genome-wide association studies (GWAS) has provided a powerful tool for the identification of common variants associated with disease, and next-generation sequencing (NGS) methods have proven similarly effective in mapping mutations underlying single gene diseases. Multiple GWAS and NGS studies have collectively identified a number of novel genetic variants associated with increased risk of ALS development. These approaches, that make assumptions

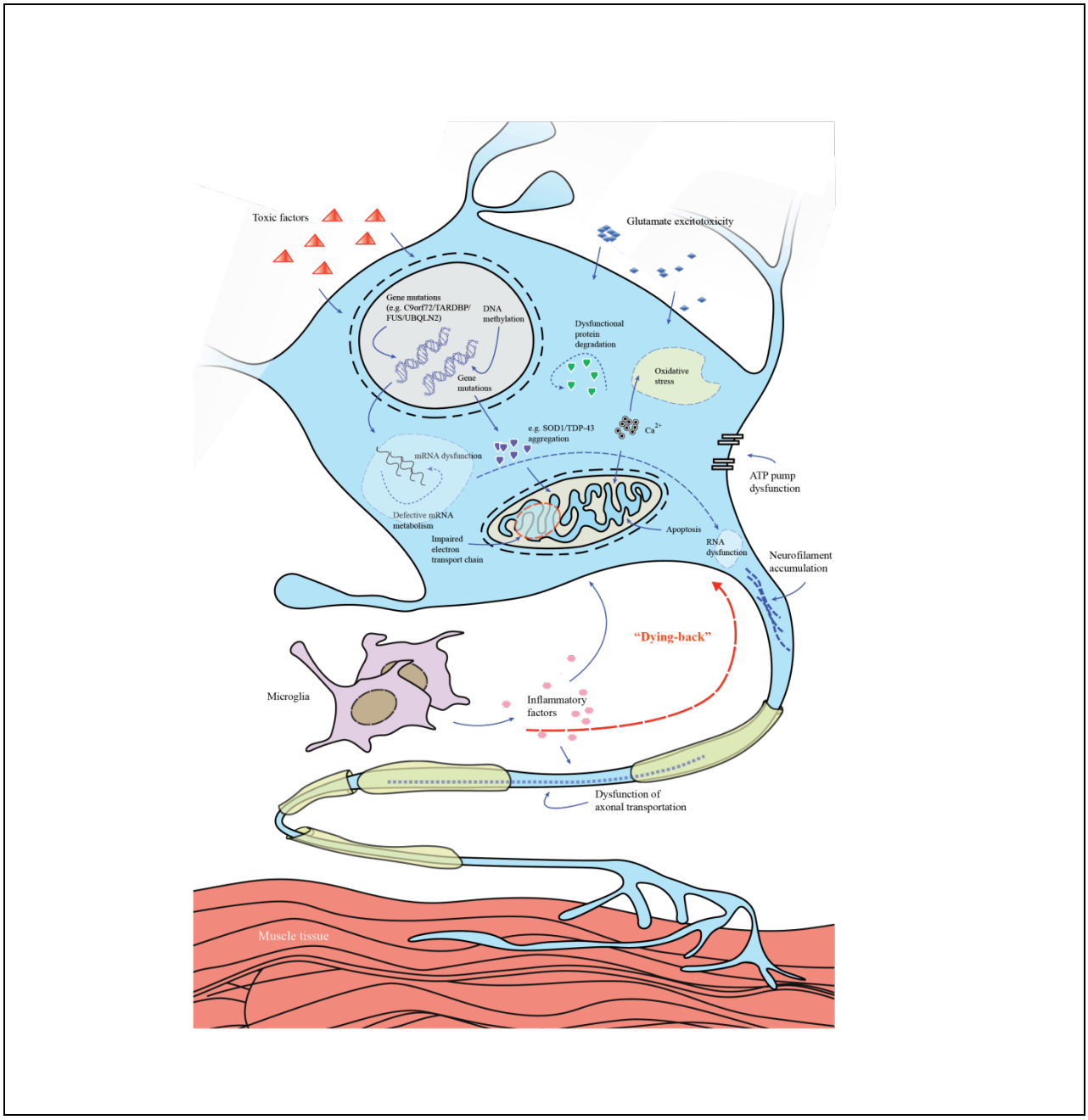

Figure 2. Proposed molecular targets and mechanisms underlying neurodegeneration in amyotrophic lateral sclerosis (ALS). Many of the initial pathological changes in models of ALS occur in the peripheral motor system, supporting a "dying-back" view of pathogenesis, though a causal primacy of lower motor neuron over upper motor neuron degeneration remains an issue of debate. The transgenic *SOD1* mouse model has been used extensively to dissect the likely pathogenic mechanisms. Many of these illustrated pathways are mechanisms of cell death common to a range of neurological disorders whereas more recent genetic discoveries have yet to be elucidated at a molecular level. Pathophysiological mechanisms involved in ALS might include combinations of glutamate excitoxicity, generation of free radicals, mutant enzymes, as well as disruption of axonal transport processes and mitochondrial dysfunction. Mutations in several ALS causative genes are related to the formation of intracellular aggregates. Mitochondrial dysfunction, which is associated with increased production of reactive oxygen species and aggregates of SOD1, might induce increased susceptibility to glutamate-mediated excitotoxicity, disturbance in energy production and apoptosis. Activation of microglia results in secretion of cytokines resulting in further toxicity.

Table 1. Causative Genes in Amyotrophic Lateral Sclerosis (ALS).

Gene	Location	Inheritance	Allelic Disorders (OMIM#)	Encoded Protein function	Ref.	Replications
*SOD1**	21q22.1	AD or AR	—	Superoxide dismutase	Rosen and others (1993)	1
*C9ORF72**	9p21.2	AD	ALS and/or FTD	Unknown	DeJesus-Hernandez and others (2011), Renton and others (2011)	1
*FUS**	16p11.2	AD or AR; de novo	ALS with or without FTD, ETM4 (614782), translocation fusion genes associated with tumors	RRM protein, RNA metabolism	Kwiatkowski and others (2009), Vance and others (2009)	1
*OPTN***	10p15-14	AD or AR	POAG (137760)	Vesicular trafficking, signal transduction	Maruyama and others (2010)	1
*SETX**	9q34	AD (juvenile onset)	SCAR1 (606002)	DNA/RNA helicase; transcription termination	Chen and others (2004)	1
*UBQLN2**	Xp11.21	XD	ALS with or without FTD	Ubiquilin protein family, protein degradation and autophagy	Deng and others (2011)	1
*ALSIN**	2q33	AR (juvenile onset)	PLSJ (606353), IAHSP (607225)	Guanine nucleotide exchange factor; AMPA receptor trafficking, endosome/membrane trafficking	Hadano and others (2001), Yang and others (2001)	1
*VAPB**	20q13.33	AD	SMAFK (182980)	ER organization	Nishimura and others (2004)	1
*SPG11**	15q14 15q15.1-	AR (juvenile onset)	SPG	Spatacsin, intracellular trafficking, motor neuron and axon development	Orlacchio and others (2010)	1
*TARDBP***	1p36	AD	ALS with or without FTD, FTLD	RRM protein, RNA metabolism	Sreedharan and others (2008)	2
*DCTN1***	2p13	AD	HMN7B (607641), Perry syndrome (168605)	Dynactin; axonal transport of vesicles and organelles	Munch and others (2004)	2
*SQSTM1***	5q35.3	AD	PDB	Autophagy, ubiquitin proteasome pathway	Fecto and others (2011)	2 (No segregation was shown)
*FIG4***	6q21	AD or AR	CMT4J (611228)	Phosphatidinositol 3,5-biphosphate 5-phosphatase; membrane trafficking, endolysosome function	Chow and others (2009)	2
*ANG***	14q11	AD	—	Ribonuclease A superfamily; RNA functions	Greenway and others (2006)	2
*DAO***	12q24	AD	—	Oxidative deamination of D-amino acids	Mitchell and others (2010)	2
*VCP****	9p13.3	AD	ALS with or without FTD, MSP (167320)	AAA+ family; endocytosis and vesicular trafficking	Johnson and others (2010)	3
*PFN1****	17p13.2	AD	—	Inhibits actin polymerization	Wu and others (2012)	3
*CHMP2b***	3p11.2	AD	FTD3	Chromatin modifying/CHMP family, endocytosis and vesicular trafficking, protein degradation	Parkinson and others (2006)	4
*PRPH***	12q12	AD	—	Type III intermediate filament protein;	Gros-Louis and others (2004),	5

				Axonal regrowth	Leung and others (2004)	
*TAF15***	17q12	AD	Translocation fusion genes in EMC	RRM protein, RNA metabolism	Ticozzi and others (2011)	5
*EWSR1***	22q12.2	AD	Translocation fusion genes in ESFT, PNE (612219), EMC	RRM protein, RNA metabolism	Couthouis and others (2012)	5
*HNRNPA1****	12q13.1	AD	MSP	RRM protein, RNA metabolism	Kim and others (2013)	6
*SIGMAR1**	9p13.3	AD (juvenile onset)	—	ER chaperone	Al-Saif and others (2011)	6
ATXN2	12q24		SCA2 (183090)	Polyglutamine protein, RNA metabolism	Elden and others (2010)	7
*NEFH***	22q12.2	AD	—	Neurofilament, heavy polypeptide	Renton and others (2011)	7
SMN1	5q13.2		SMA (253300)	RRM protein, RNA metabolism	Corcia and others (2002), Blauw and others (2012)	7

Mode types: AD = autosomal dominant; AR = autosomal recessive; XD = X-linked dominant. *, **, and *** indicate the gene was first identified in ALS patients by positional cloning, candidate gene screening, and whole exome sequencing (WES), respectively. SCA2 = spinocerebellar ataxia-2; CMT4J = Charcot–Marie–Tooth, autosomal recessive, type 4J; PLSJ = primary lateral sclerosis, juvenile; IAHSP = infantile onset ascending spastic paralysis; SCAR1 = spinocerebellar ataxia, autosomal recessive; ETM4 = hereditary essential tremor-4; SMAFK = spinal muscular atrophy, late-onset, Finkel type; POAG = primary open angle glaucoma; MSP = multisystem proteinopathy; FTD3 = frontotemporal dementia, chromosome 3-linked; EMC = extraskeletal myxoid chondrosarcoma; ESFT; Ewing sarcoma family of tumors; PNE = peripheral neuroepithelioma; PDB = Pagets disease of bone; SPG = spastic paraplegia. Replication: 1. Linkage in at least one family (logarithm of odds [LOD] >3). The mutations/s segregates with disease, is not found in normal samples, and mutations were subsequently found in sALS or other small families; 2. Variants were found to segregate in multiple small families (but were too small to achieve genome-wide significance in linkage analysis) as well as found in multiple sporadic cases and are not found in normal samples; 3. Gene contains variants that segregate in multiple small families but have not been reported in sporadic cases, and are not found in normal samples; 4. Variants described in one small family and multiple sporadic cases and are not found in normal samples. Functional tests assessing the affect of the mutation on the protein function support a role in ALS; 5. Mutations have been reported in multiple sporadic cases and not in normal samples. Functional tests assessing the affect of the mutation on the protein function support a role in ALS; 6. Rare variants have been found in the gene in only one family; 7. copy number variants (CNV)/particular alleles present in ALS patients significantly more than control.

about the underlying genetic architecture of disease especially for sporadic cases, can be complex to perform (Hardiman and others 2007), and depend on accurate clinical phenotyping for success. Thus, whilst many robust associations have been made with rare mutations and common variants in ALS, false positive findings have also occurred (Garber 2008). Given this context it is important to examine the performance and limitations of GWAS along with the expectations for NGS in the study of ALS.

Are Sporadic and Familial ALS Different?

The genetic architecture of common diseases can be investigated by studying the recurrence rate in relatives of affected individuals (Risch 1987). A twin study has shown that for ALS the concordance rate in identical monozygotic twins is ~10%, whereas the concordance rate in dizygotic twins is very low (0/122 twins) as is the parent-offspring concordance rate (~1%) (Al-Chalabi and others 2010). From the twin concordance values, the heritability of ALS has been calculated to be 61%, taking into account the low population prevalence of the disease (Al-Chalabi and others 2010). For monogenic autosomal diseases, the parent–offspring rate is generally high unless the penetrance is low (Risch 1990a, 1990b). Thus the low parent–offspring and twin concordance rates are most consistent with models involving either multiple variants, either common or rare, especially variants with low penetrance, contributing to each case, rather than single highly penetrant variants in each individual.

The boundary between sporadic and familial ALS is becoming far more blurred than their nomenclature suggests. Although it has been reported that 10% of ALS is familial, when the genealogy of sALS cases is investigated more thoroughly, the prevalence of a family history increases to 20% in some prospective studies (Chio and others 2011; van Es and others 2010). Establishing familiality is difficult in late onset diseases, and it is likely that this leads to underestimation of the familiality of ALS. It is often quite difficult for adult children of ALS patients to accurately recall their parent's symptoms, particularly if decades have passed since the event, which occurred when the diagnosis of ALS may not have been considered.

The idea that sALS and fALS may share a diverse genetic aetiology is consistent with the observation that sALS and fALS are clinically indistinguishable (Andersen and others 1997), and because of increasing demonstration that "sporadic" cases are often caused by the same mutations that have previously been reported in familial cases. Familial and sporadic cases share histopathological characteristics; for example, superoxide dismutase 1 (SOD1)–positive and TAR DNA-binding protein 43 (TDP-43)–positive inclusions have been found in both sALS and fALS (Bosco and others 2010; Forsberg and others 2010). Moreover, first-

degree relatives of patients with ALS have an increased incidence compared with unrelated individuals (Fallis and others 2009). Additionally, the hexanucleotide repeat expansion mutation in the *C9orf72* gene is associated with a founder haplotype. The expansion is found both in familial cases and in a high proportion of "sporadic" ALS cases (21% of Finnish sALS) (Renton and others 2011), and, albeit at low frequency, in control populations (prevalence of expansions 0.15% in 1958 British Birth Cohort) (Beck and others 2013). The *C9orf72* expansion occurs on a common risk haplotype, which is prone to expansion of the hexanucleotide repeat for reasons that are not understood. Thus whilst the common founder haplotype is prevalent in the community, further mutation (expansion of the hexanucleotide repeat) on this haplotype may lead to ALS, and would at least partially explain the occurrence of sALS due to *C9orf72* expansions in families with no previous history of the disease (DeJesus-Hernandez and others 2011). The alternative is that the repeat expansion is not fully penetrant and modifier loci or environmental factors are also required for *C9orf72* repeat expansions associated ALS pathogenesis.

Combined, this evidence indicates that sALS diagnosis does not exclude the possibility of familial inheritance even with "definite" negative familial histories. The term "sporadic" has also been used for other neurological disorders such as autism, in which a high degree of genetic heterogeneity has now been uncovered (O'Roak and others 2012). We agree with the suggestion that the term "sporadic" is inappropriate as it suggests only non-genetic causes, whereas the term "isolated" ALS might be more accurate as it encompasses environmental and/or genetic causes.

Prevalent Heterogeneity in Amyotrophic Lateral Sclerosis

Amyotrophic lateral sclerosis demonstrates both clinical and genetic heterogeneity, but to date genetic explanations for the clinical heterogeneity have been incomplete.

Three clinical variants are widely recognized, including progressive bulbar palsy (PBP), classic onset and progressive muscular atrophy (PMA) (del Aguila and others 2003). By and large, bulbar onset appears to advance more rapidly and PMA may have a better prognosis (del Aguila and others 2003; Norris and others 1993). Besides these three, other forms have been recognized in recent years. Brachial amyotrophic diplegia (flail arm syndrome, FA) and the pseudopolyneuritic variant (flail leg syndrome, FL) are instructive examples. Compared with classic type, FA and FL have a prolonged course (Wijesekera and others 2009). Furthermore, these clinical variants appear to have a male predominance, which may suggest that other gender-related genes are involved, although such genes are yet to be identified (Gamez and others 1999).

Recent genetic studies have taken clinical heterogeneity into account, but have yet to identify individual variants or genomic profiles associated with different ALS clinical patterns (Dunckley and others 2007; Kwee and others 2012). An Italian study (Orsetti and others 2011) involving 228 ALS patients reported a potential correlation between the upper motor neuron-predominant phenotype and *KIFAP3*. However, these findings have not been replicated in subsequent studies (Table 3). This may be because in some studies the clinical subgroups were not clearly defined in the same manner as the Italian study, and in others the study power was likely insufficient. Whatever the explanation, clear genotype–phenotype correlation to explain the different presentations is lacking.

Studies of ALS have been particularly successful in identifying disease-causative mutations (Table 1). Clinical heterogeneity is seen in unrelated ALS cases that share the same mutation, but also in carriers from within the same family. On the other hand families with mutations in different genes are often clinically indistinguishable. In addition to the locus heterogeneity, there is also extensive allelic heterogeneity. For example, at least 160 disease-causative mutations in *SOD1* have been reported (Birve and others 2010; Zinman and others 2009). If genetic or locus heterogeneity is as extensive in sporadic ALS, genetic studies will struggle to identify the genes involved, unless quite large case and control cohorts are studied.

Amyotrophic Lateral Sclerosis Studies in Different Ethnic Populations

Genetic differences between ethnic groups are often a nuisance in genetic studies. Nonetheless, they can be useful tools where the variation in linkage disequilibrium and polymorphism frequencies between populations provides useful contrasts that can be informative in identifying key disease-associated variants. There are no apparent differences in the incidence or prevalence of ALS between ethnic groups. However, several spatial clusters of ALS have been reported that may be explained by a higher frequency of particular mutations in geographical regions. A spatial cluster of ALS in south-eastern Finland (Sabel and others 2003) is likely explained by the high occurrence of the *C9orf72* hexanucleotide repeat in ALS cases from Finland. Similarly, although the *C9orf72* repeat expansion is generally rare in Japanese ALS cases, it has been reported in cases from the Kii peninsula where another cluster had also been observed (Yoshida and others 1998). Perhaps of more interest for disease progression is the *SOD1* mutation (D90A) that has been

shown to be responsible for slowly progressive, autosomal recessive ALS in Sweden and Finland (Andersen and others 1997). A total of 1% to 2.5% of the population are heterozygous carriers of this mutation, yet do not develop ALS. However, in other geographical locations, heterozygous carriers of the same mutation develop classical ALS (Robberecht and others 1996). Whether this is caused by environmental effects or modifier genetic effects is unknown. Thus, there is likely to be genetic heterogeneity between ethnic groups, providing new information about the genetic causes of a disease. To date nearly all large-scale genetic studies in ALS have been performed in populations of European descent, with only limited studies performed in Asians, and none that the authors are aware of in African Americans. Such studies would be beneficial in elucidating the genetic contribution to ALS globally.

Genome-wide Association Studies in Amyotrophic Lateral Studies

Traditional genetic strategies such as linkage studies or candidate gene sequencing have had major success in finding mutations in familial ALS. However, their potential is limited by the low availability of families with recurrent disease, challenges associated with family studies in late onset diseases, and by the low power of linkage analysis to identify low-penetrant variants. As cases are often misdiagnosed and/or have died before recurrence occurs within individual families, few families with DNA samples from multiple affected and unaffected members from multiple generations are available. The variability in the age of onset, even within families, means that coding unaffected members of families can be misleading, when they may develop ALS later in life. Additionally there are now several examples of ALS mutations with incomplete penetrance (for example in *SOD1* (Gamez and others 2011) and *TARDBP* (Orru and others 2012)), the presence of which can also affect linkage studies. Given these challenges with traditional linkage mapping approaches, and the possibility that many sALS cases may be because of low penetrance, possibly common founder variants, in recent years alternate methods involving GWAS and exome sequencing have been employed to identify the genetic variants.

The Design of Amyotrophic Lateral Sclerosis GWAS

Genome-wide association studies identify loci associated with disease where cases or controls share stretches of DNA (haplotypes) in which a disease-causing mutation has occurred in a distant common founder. By typing SNPs across the genome, these disease-associated haplotypes are identified by differences in frequencies of SNPs between cases and controls because of the marker SNPs being found on the same haplotype as the disease-causing variant, a feature termed 'linkage disequilibrium'. Association in GWAS studies may thus represent true association (where the genotyped, associated SNP is directly involved in the disease), linkage disequilibrium (where the genotyped, associated SNP is merely a marker for a haplotype bearing the true disease-causing variant), population stratification (where differences in ethnicity between cases and controls lead to spurious associations), technical artifacts in the genotyping (such as batch effects, differential missingness), or cryptic relatedness (where the assumption that cases and controls are not closely related to one another is not met, leading to inflation of association test statistics). Careful quality control checks of GWAS data have been developed, and if these are used then findings that achieve genome-wide significance (generally thought to be $P < 5 \times 10^{-8}$, although this does depend on the prior probability of a true positive finding) have generally proven robust and reproducible.

Sample size requirements for GWAS are large and challenging to achieve, particularly for low frequency diseases such as ALS (Purcell and others 2003) (Figure 3). For most diseases, this has led to the formation of large consortia pooling resources to achieve adequate power. Where separate studies are performed, meta-analysis methods and software for either the original genotype data or summary statistics have been developed to achieve sufficient power. There has been some success employing GWAS in ALS (Table 2), and some of that raw data has been made publicly available to augment resource for other studies of ALS. The largest GWAS to date in ALS involved 6100 cases and 7125 controls (Fogh and others 2014). Although this is evidently a large study, the sample size requirement presented in Figure 1 suggest that even larger studies will be required to identify further loci, particularly to identify low frequency variants, even if they have quite high penetrance. For example, to achieve 80% power to identify a variant with minor allele frequency of 1% and odds ratio for disease of 5 will require >10,000 cases to be studied! Such sample sizes have been achieved for diseases less common than ALS. For example, in Crohn's disease, which has a prevalence of 0.1%, more than 77,000 cases have been reported in a single study (Jostins and others 2012).

MAF=0.25,p value=5x10^{-8}	
OR	Number of cases in 80% power
1.1	41,625
1.2	11,473
1.3	5,593
1.4	3,436
1.5	2,392
1.6	1,800
1.7	1,428
1.8	1,178
1.9	999
2	867
3	386
4	267
5	215
6	187
7	169
8	157
9	148
10	141
MAF=0.01,p value=5x10^{-8}	
OR	Number of cases in 80% power
1.1	13,533,991
2	141,432
3	37,083
4	17,264
5	10,160
6	6,795
7	4,926
8	3,774
9	3,010
10	2,476
20	795
30	460
40	330
50	262
60	222
70	195
80	176
90	162
100	151

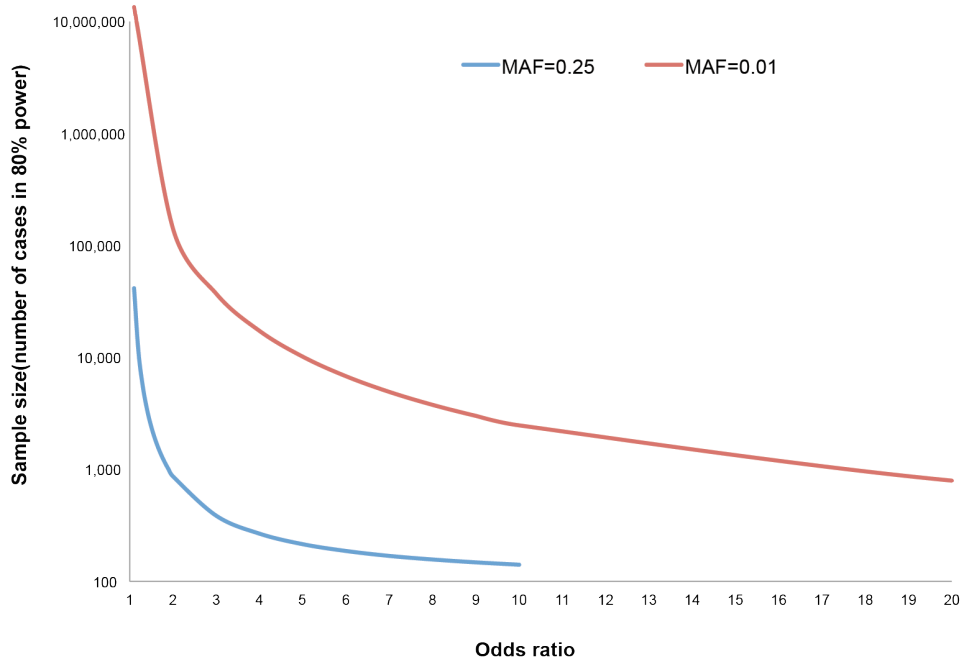

Figure 3. Genetic power calculation for the sample size according to the different minor allele frequencies (MAFs).

Table 2. Main Results Reported by Genome-wide or Large-scale Strategies in ALS Studies.

						Original Study				
						MAF			Pooled *P* Value	
No.	Ref.	SNP	Chr	Gene	Description	Controls	Patients	Pooled OR (95% CI)	*P* Value	Bonferroni Correction
1	van Es and others (2007)	rs2306677	12	*ITPR2*	A calcium channel on the endoplasmic reticulum	0.07	0.11	1.58 (1.30-1.91)	3.28×10^{-6}	Exceed
2	Dunckley and others (2007)	rs6700125	1	*FLJ10986 (FGGY)*	Unknown, possible role in metabolism	0.32	0.41	1.35 (1.13-1.62)	6×10^{-4}	Exceed
3	Cronin and others (2008)	rs10260404	7	*DPP6*	A transmembrane protein binding A-type neuronal potassium channels	0.34 (Irish)	0.42 (Irish)	1.37 (1.2-1.56)	2.53×10^{-6}	Failed
4	van Es and others (2009)	rs12608932	19	*UNC13A*	Encoding presynaptic proteins found in central and neuromuscular synapses	0.37	0.40	1.2	2.50×10^{-14}	Exceed
5		rs2814707	9	*MOBKL2B*	Unknown	0.23	0.26	1.16	7.45×10^{-9}	Exceed
6		rs3849942	9	*C9orf72*	Unknown	0.23	0.26	1.15	1.01×10^{-8}	Exceed
7	Chio and others (2009)	rs2708909	7	*SUNC1*	Encodes a 40.5-kDa nuclear envelope protein Sad1 and UNC84	0.45	0.50	1.17 (1.11–1.23)	6.98×10^{-7}	Failed
8		rs2708851	7	*SUNC1*		0.45	0.50	1.17 (1.11-1.23)	1.16×10^{-6}	Failed

					domain containing 1					
9	Iida and others (2011)	rs2275294	20	*ZNF512B*	A regulator of the TGFβ signaling pathway	0.41	0.48	1.32 (1.21-1.44)	6.70×10^{-10}	Exceed
10	Deng and others (2011)	rs6703183	1	*CAMK1G*	It encodes Ca^{2+}/calmodulin-dependent protein kinase which belonging to the CaM kinase family.	0.34 (combined)	0.41 (combined)	1.31 (combined)	*P* combined = 2.92×10^{-8}	Failed
11		rs8141797	22	*SUSD2*	SUSD2 is mainly expressed in the brain, kidney and lung and codes for a novel tumor suppressor	0.10 (combined)	0.15 (combined)	1.52 (combined)	*P* combined = 2.35×10^{-9}	Failed
12	Fogh and others (2014)	rs34517613	17	Unknown	Unknown	0.13	0.11	0.82, (0.76-0.87)	1.11×10^{-8}	Failed

ALS = amyotrophic lateral sclerosis; Chr = chromosome; CI, confidence interval; MAF: minor allele frequency; OR = odds ratio; Ref. = reference; SNP = single nucleotide polymorphisms.

As discussed above there are unique challenges in recruiting patients with ALS that make it particularly difficult to achieve such large numbers, but given the great potential of genetics in this disease, greater emphasis on recruitment is required, and further collaborative studies encouraged. In addition, large GWAS meta-analytical studies including both Caucasian and non-Caucasian populations are likely to be informative.

The Results from GWAS

In the past 6 years, a conservative estimate of more than 8000 ALS patients (including patients of white European and Asian descent), and many more control subjects, have been involved in large scale ALS studies. This has resulted in the identification of SNPs that appear to be associated with the disorder. The discovery of association with the *C9orf72* region is the most well known of these. In 2010, three ALS-GWASs reported a novel susceptibility locus associated with the disease on chromosome 9p where *C9orf72* is located (van Es and others 2009; Laaksovirta and others 2010; Shatunov and others 2010). Another independent GWAS of ALS with pathologically confirmed FTLD/TDP (frontotemporal lobar degeneration/TAR DNA-binding protein) (Gijselinck and others 2010) identified the same chromosomal region. Linkage of ALS–FTLD to this region was first identified in 2006 (Morita and others 2006; Vance and others 2006), indicating that the chromosome 9p gene defect overlaps the candidate region for both FTD and ALS. Three genes were found in a haplotype block with strong association: *MOBKL2B*, *IFNK*, and *C9orf72*. Since this time, the hexanucleotide (GGGGCC)n repeat expansion in *C9orf72* has been identified as the gene defect associated with ALS and FTD (DeJesus-Hernandez and others 2011; Renton and others 2011). These correlations have been reported in many autosomal dominantly inherited cases with both FTD and ALS (Gijselinck and others 2010; Herdewyn and others 2012), and recently *C9orf72* expansions have also been found in cohorts with a variety of neurodegenerative diseases including Alzheimer disease, sporadic Creutzfeldt-Jakob disease, Huntington disease-like syndrome, and other non-specific neurodegenerative disease syndromes in addition to FTD and ALS (Beck and others 2013).

The above experimental conclusions become even more convincing when the data are subjected to meta-analysis and pooled analysis. The need for such approaches is highlighted by the high rate of non-replication of findings between ALS GWAS studies

Table 3. Major Replications (Including GWAS) for Variants in ALS Studies.

			Gene Replication					
Ref.	Region of Cohorts	Study Method	*ITPR2*	*FGGY*	*DPP6*	*KIFAP3*	*UNC13A*	Chromosome 9p21 Locus
Cronin and others (2008)	Ireland	GWAS	×	×	$P = 2.53 \times 10^{-6}$	—		
van Es and others (2008)	USA/Dutch	Replication	—		$P = 5.04 \times 10^{-8}$	—		
van Es and others (2009a)	North Europe	Replication	—	×	—			
Daoud and others (2010)	France and Quebec	Replication	—	×	×	—		
Landers and others (2009)	Europe/USA	GWAS	×	×	×	Corrected P = .021	—	
Chio and others (2009)	Italy/Germany/USA	GWAS	×	×	×	—		
Cronin and others (2009)	Irish/Polish/USA/Dutch	GWAS	—		×	—		
van Es and others (2009b)	Europe/USA	GWAS	×	×	×	—	Table 2	
Shatunov and others (2010)	Europe/USA	GWAS	×	×	×	—	×	$P = 5{\cdot}14 \times 10^{-8}$
Laaksovirta and others (2010)	Finland	GWAS	—					×
Fogh and others (2011)	Italy	Replication	—	×	—			
Orsetti and others (2011)	Italy	Replication	—			×	—	
Iida and others (2011b)	Japan	Large-scale association study	—					
Iida and others (2011a)	China/Japan	Replication	—				×	×
Kwee and others (2012)	USA	High-density GWAS	×	×	×	×	×	×
Chen and others (2012)	China	Replication	×	×	×	×	—	
Ratti and others (2012)	Italy	Replication	—					$P = 8 \times 10^{-9}$; OR = 4.0
Diekstra and others (2012)	Dutch/Belgian/Swedish	Replication	—				$P <$ .005 (sic)	—
Mok and others (2012)	Europe/USA	Meta-analysis	—					Overall P = .0088
Deng and others (2011)	China	GWAS	×	×	×	×	×	×

GWAS = genome-wide association studies. Ref. = reference; —, the direct replication for the implicated gene was not included; ×, negative results.

(Table 3), which very likely represent type 1 errors due to insufficient sample size and statistical power. This is a particular problem in genetic studies of ALS, likely because the disease is rare and it is hard for individual groups to achieve adequate sample sizes. Indeed with the exception of the chromosome 9p21 locus, most associated SNPs or genes show poor replication in different populations. This is also an issue in transethnic studies, where for example even the findings for the chromosome 9p21.2 locus are inconclusive in Asian populations probably because the repeat expansion is rare. Robustly proven differences in associations at specific loci can be highly informative regarding the primary disease-causative variants, which can be hard to pinpoint in specific populations, particularly at loci with extensive linkage disequilibrium surrounding them.

Copy Number and Structural Variants in Amyotrophic Lateral Sclerosis

Microarray SNPs used in GWAS have modest capability to genotype copy number variants (CNVs) and other structural variants. When inherited, these are typically in linkage disequilibrium with SNPs and at most common loci have been detected by GWAS even if the CNV has not been typed itself (Wellcome Trust Case Control

Consortium and others 2010). For ALS only a handful of CNV analyses have been reported. The most well studied CNV is of the *SMN1* gene in which loss of function mutations cause autosomal recessive spinal muscular atrophy. Several reports suggested that duplication of the *SMN1* gene is found at a higher frequency in ALS patients (Corcia and others 2002; Corcia and others 2006; Veldink and others 2005) and an association of ALS with *SMN1* copy number was later confirmed in an analysis of 3500 ALS cases (Blauw and others 2012). Aside from the *SMN1*, other CNV analyses in ALS have not had sufficient power to confirm correlations of rare events (Cronin and others 2008). The lack of CNV associations with ALS are likely because of limits of current methods for genotyping CNVs, or that they are rare rather than that they do not occur. Recently, whole genome sequencing successfully uncovered de novo CNVs in subjects with intellectual disability (Gilissen and others 2014), and a similar approach may be required to uncover ALS CNVs, if they exist.

Genome-wide Association Studies and Rare or De Novo Variants

Given the limited sample sizes studied and typically weaker linkage disequilibrium around rare and low frequency variants, association signals from such variants are frequently missed by GWAS. Of particular relevance to ALS, GWAS is not capable of detecting de novo sporadic mutations that may be responsible for a significant fraction of cases, as opposed to variants arising in a distant common founder.

Thus while GWAS is a powerful hypothesis-free approach to identifying genetic variants in common human diseases, it is not capable of identifying all variants, including a range of variants already known to be involved in ALS.

Sequencing for Amyotrophic Lateral Sclerosis Gene Mapping

As GWAS studies are poor at identifying low frequency (minor allele frequency [MAF] <5%) and rare (MAF <1%) alleles, it has been suggested that large numbers of such variants may explain the failure of GWAS to identify the variants responsible for a large proportion of the heritability of individual diseases (Manolio and others 2009). This remains however an unproven hypothesis, with little evidence to date to support the existence of multiple low frequency or rare variants with greater penetrance than common variants. The few reported examples of highly penetrant rare alleles have generally occurred in genes already known to be associated with disease through common variant studies, such as *IL23R* (Momozawa and others 2011), *CARD9* (Rivas and others 2011), and *IFIH1* (Nejentsev and others 2009). This may represent experimental bias as sufficiently large studies to detect rare variants are still works in progress. In ALS though, it is clear that rare highly and lowly penetrant variants do contribute significantly to the disease—these are the variants that cause fALS, and are increasingly being identified in sALS.

Next-generation sequencing techniques, including whole exome sequencing (WES) and whole genome sequencing (WGS) represent a powerful new paradigm with regard to addressing monogenic disorders. These techniques have enabled gene-mapping studies for rare genetic diseases to be performed in the absence of multigeneration family material. Instead, either small numbers of unrelated affected individuals or small families can now be successfully investigated. This has proven revolutionary in monogenic disease research, stimulating another boom in gene discoveries for diseases that had previously not been mappable because of insufficient family material. The approach has already proven productive in fALS, where it has been used to identify disease-causing mutations in *VCP* and *PFN1* (Johnson and others 2010; Wu and others 2012) that are strongly supported as ALS-causing mutations from functional studies. Additionally, a WES study using samples from 47 sporadic ALS cases and their unaffected parents, have identified de novo mutations in several candidate genes, including as *SS18L1*, encoding proteins that regulate chromatin (Chesi and others 2013).

The same approaches using NGS to map monogenic diseases are now being used to map rare high penetrant variants in more genetically complex diseases, such as sALS. Unlike monogenic diseases, ALS creates a new challenge because of the genetic and locus heterogeneity and the possibility that ALS is oligo- or polygenic, with more than one disease-causing variant operating in individuals. This requires much larger sample sizes to be studied and more complex analyses to be performed, inevitably involving greater cost. However, given that ALS may be substantially caused by rare or even de novo mutations, these challenges will have to be faced to advance our understanding of the genetics of this disease. In the past arguments for variants being disease-causing have been strengthened by evidence of segregation with disease in fALS, along with identification of mutations in sporadic cases. For *PRPH* and *EWSR1,* mutations have only been described in sALS; however, convincing functional data accompanied the variant discovery. Functional methods for determining the relevance of rare sequence variants identified by NGS to ALS pathogenesis will continue to be necessary in the study of genetic causes of sALS.

The diagnostic benefits that NGS will provide for ALS patients and their neurologists are obvious. Currently the average time from symptom onset to

confirmed diagnosis for ALS patients is nearly 1 year, a time in which misdiagnosis and sometimes even invasive surgery may occur (Paganoni and others 2014). Early diagnosis will be vital for clinical trials where early treatment before significant damage has occurred will be of utmost importance. Early diagnosis may be facilitated by sequencing early symptomatic cases, or by screening in families known to carry high penetrant variants. Furthermore, knowledge of the genetic cause of ALS in patients participating in clinical trial will informative, where in the context of a genetically heterogeneous disease success or failure of treatments may depend on the original genetic cause. Until treatments reach the clinic, the multitude of researchers using donated material from ALS patients to try to understand why motor neurons degenerate, will benefit by being able to group samples based on genetics to more meaningfully interpret data generated.

Limitations of Next-Generation Sequencing in Amyotrophic Lateral Sclerosis

Although there are clear advantages to NGS, it is not without drawbacks. As the name implies, WES is designed to capture the whole exome. In reality it does not, and about 5% of the exome is not captured in typical exome sequencing experiments. Furthermore, not all monogenic diseases are caused by mutations in coding regions, and therefore may not be detected using exome sequencing. An additional technical challenge is that WES using short-read sequencing technologies such as Illumina sequencing chemistry has difficulties in sequencing repeats and insertion/deletions, as the short reads cannot be unambiguously aligned against reference genomes. Some relevant examples of the limitations of WES exist in ALS gene-mapping include the following:

- TDP-43 and FUS are known to regulate splicing of pre-mRNA by binding to sites within introns (Arnold and others 2013; Lagier-Tourenne and others 2012; Polymenidou and others 2011); WES may not identify mutations at these binding sites and thus if mutations at sites in genes targeted by TDP-43 or FUS were involved in ALS, they would not be detected by WES.
- An intronic hexanucleotide repeat in *C9orf72* gene has been identified though linkage mapping and several GWASs and the finding has been replicated independently. This locus is difficult to detect by WES as it is intronic, and also by WGS, because of the difficulty aligning short-read data of stretches of microsatellite and minisatellite DNA sequences. This weakness may be overcome using longer read technologies such as single molecule real-time sequencing (Pacific Bioscience sequencing), which has a read length of up to 10 kb, compared with short 2 × 100- or 2 × 150-bp paired-end reads typically produced by, for example, Illumina sequencing.

Thus, sequencing is not at this point a technology capable of identifying all potential ALS-causative variants, and it should be noted still requires the large sample sizes discussed above to achieve adequate power. NGS approaches have great potential and are capable of identifying a substantial fraction of variants that are not within the scope of GWAS. Advances in sequencing technologies, decreasing cost of sequencing and initiatives such as Project MinE (www.projectmine.com) aiming to raise funds and generate whole genome sequence from 15,000 ALS subjects, will lead to data being rapidly produced over the coming years, allowing in-depth genetic analysis.

Genome-wide Methylation Analysis

Several different methods involving either microarrays or DNA sequencing now enable genome-wide characterization of DNA methylation.

Epigenetic variation including DNA methylation is a major determinant of gene transcription. Improved methods to assess DNA methylation in particular has led to increasing interest in the role of epigenetic variation in human disease, including ALS (Figueroa-Romero and others 2012). An early microarray based analysis of methylated DNA immunoprecipitated from post-mortem brain found differential gene methylation in ALS cases (Morahan and others 2009) and a second using ALS spinal cord reported changes in global methylation and hydroxymethylation (Figueroa-Romero and others 2012). These studies represent an initial identification of epigenetic regulatory changes in ALS, however, were only performed on 10 and 12 ALS samples, respectively. Larger studies with more sensitive and robust methods are clearly required to uncover the contribution of epigenetics to ALS.

Aside from genome-wide methylation analyses, several gene specific analyses of CpG methylation within promoters of ALS genes and candidate genes have been performed. No difference in DNA methylation was observed in ALS cases in the promoter regions of *SOD1*, *VEGF* (Oates and Pamphlett 2007) and *EAAT2* (Yang and others 2010). However, CpGs upstream of the *C9orf72* repeat expansion in some expansion carriers were hypermethylated and the corresponding mRNA was down-regulated in a limited number of samples tested(Xi and others 2013). Methylation of upstream CpG islands is not novel in terms of repeat expansion mutations and the role that methylation at the *C9orf72* locus plays in ALS

pathogenesis is yet to be determined and requires further analysis in larger cohorts of carriers.

Other Methods Combined with Whole Genome Strategies

Commonly, alleles showing a correlation between allele load and expression levels are defined as expression quantitative trait loci (eQTL). Diekstra and others (2012) conducted a two-stage genome-wide screen for eQTLs associated with ALS. Data from a two-stage GWAS were combined to prioritize eQTLs identified in the first stage. While identifying candidate genes for sporadic ALS, most notably *CYP27A1*, the study demonstrated the potential of an integrated approach in identifying causative genetic variants in ALS.

Both methylation and expression analyses in ALS suffer from issues related to tissue sampling because the site of disease pathology is not amenable during the life of the patients. Postmortem material, although valuable, represents end-stage of disease when significant damage and neuron loss has already occurred. It is also difficult to obtain the large numbers of post-mortem samples required for genome-wide analyses. Other sources of RNA and DNA that are more suitable for the large studies required for such a genetically and clinically heterogenous disorder as ALS will need to be considered.

Expectation for Future Amyotrophic Lateral Sclerosis Genetic Studies

Genetic discoveries in ALS have provided major advances in our understanding of the diseases causation, and the development of new methods for gene mapping particularly using high-throughput sequencing promises much. There is also a clear need to perform further studies in populations of different ethnicity. Systems biology approaches including multiple modalities such as DNA sequencing, transcriptomics, proteomics and epigenomics are likely to be useful, particularly with functional analysis of the mechanisms by which genetic mutations cause ALS. As ever there is no "one-size-fits-all" solution to ALS research, but the increasing capability of our research tools is bringing major advances to studies of this disease. In other disease areas such as immune-mediated diseases, success in gene-mapping has been so great in the past 5 years of the GWAS era that the bottleneck has clearly moved on to determining how disease-associated variants lead to disease. The authors expect that this will similarly be the problem facing ALS research in the near future.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Ji He was funded by a grant from the National Natural Sciences Foundation of China (81030019); Marie Mangelsdorf was supported by the Ross Maclean Senior Research Fellowship; Matthew A. Brown was funded by a National Health and Medical Research Council (Australia) Senior Principal Research Fellowship.

References

Al-Chalabi A, Fang F, Hanby MF, Leigh PN, Shaw CE, Ye W, and others. 2010. An estimate of amyotrophic lateral sclerosis heritability using twin data. J Neurol Neurosurg Psychiatry 81(12):1324–6.

Al-Saif A, Al-Mohanna F, Bohlega S. 2011. A mutation in sigma-1 receptor causes juvenile amyotrophic lateral sclerosis. Ann Neurol 70(6):913–9.

Andersen PM, Nilsson P, Keranen ML, Forsgren L, Hagglund J, Karlsborg M, and others. 1997. Phenotypic heterogeneity in motor neuron disease patients with CuZn-superoxide dismutase mutations in Scandinavia. Brain 120(Pt 10):1723–37.

Arnold ES, Ling SC, Huelga SC, Lagier-Tourenne C, Polymenidou M, Ditsworth D, and others. 2013. ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. Proc Natl Acad Sci U S A 110(8):E736–45.

Beck J, Poulter M, Hensman D, Rohrer JD, Mahoney CJ, Adamson G, and others. 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(3):345–53.

Birve A, Neuwirth C, Weber M, Marklund SL, Nilsson A-C, Jonsson PA, and others. 2010. A novel SOD1 splice site mutation associated with familial ALS revealed by SOD activity analysis. Hum Mol Genet 19(21):4201–6.

Blauw HM, Barnes CP, van Vught PW, van Rheenen W, Verheul M, Cuppen E, and others. 2012. SMN1 gene duplications are associated with sporadic ALS. Neurology 78(11):776–80.

Bosco DA, Morfini G, Karabacak NM, Song Y, Gros-Louis F, Pasinelli P, and others. 2010. Wild-type and mutant SOD1 share an aberrant conformation and a common pathogenic pathway in ALS. Nat Neurosci 13(11):1396–403.

Byrne S, Walsh C, Lynch C, Bede P, Elamin M, Kenna K, and others. 2011. Rate of familial amyotrophic lateral sclerosis: a systematic review and meta-analysis. J Neurol Neurosurg Psychiatry 82(6):623–7.

Chen Y, Zeng Y, Huang R, Yang Y, Chen K, Song W, and others. 2012. No association of five candidate genetic variants with amyotrophic lateral sclerosis in a Chinese population. Neurobiol Aging 33(11):2721.e3–e5.

Chen YZ, Bennett CL, Huynh HM, Blair IP, Puls I, Irobi J, and others. 2004. DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). Am J Hum Genet 74(6):1128–35.

Chesi A, Staahl BT, Jovicic A, Couthouis J, Fasolino M, Raphael AR, and others. 2013. Exome sequencing to identify de novo mutations in sporadic ALS trios. Nat Neurosci 16(7):851–5.

Chio A, Borghero G, Pugliatti M, Ticca A, Calvo A, Moglia C, and others. 2011. Large proportion of amyotrophic lateral sclerosis cases in Sardinia due to a single founder mutation of the TARDBP gene. Arch Neurol 68(5):594–8.

Chio A, Schymick JC, Restagno G, Scholz SW, Lombardo F, Lai SL, and others. 2009. A two-stage genome-wide association study of sporadic amyotrophic lateral sclerosis. Hum Mol Genet 18(8):1524–32.

Chow CY, Landers JE, Bergren SK, Sapp PC, Grant AE, Jones JM, and others. 2009. Deleterious variants of FIG4, a phosphoinositide phosphatase, in patients with ALS. Am J Hum Genet 84(1):85–8.

Corcia P, Camu W, Halimi JM, Vourc'h P, Antar C, Vedrine S, and others. 2006. SMN1 gene, but not SMN2, is a risk factor for sporadic ALS. Neurology 67(7):1147–50.

Corcia P, Mayeux-Portas V, Khoris J, de Toffol B, Autret A, Muh JP, and others. 2002. Abnormal SMN1 gene copy number is a susceptibility factor for amyotrophic lateral sclerosis. Ann Neurol 51(2):243–6.

Couthouis J, Hart MP, Erion R, King OD, Diaz Z, Nakaya T, and others. 2012. Evaluating the role of the FUS/TLS-related gene EWSR1 in amyotrophic lateral sclerosis. Hum Mol Genet 21(13):2899–911.

Cronin S, Berger S, Ding J, Schymick JC, Washecka N, Hernandez DG, and others. 2008. A genome-wide association study of sporadic ALS in a homogenous Irish population. Hum Mol Genet 17(5):768–74.

Cronin S, Hardiman O, Traynor BJ. 2007. Ethnic variation in the incidence of ALS: a systematic review. Neurology 68(13):1002–7.

Cronin S, Tomik B, Bradley DG, Slowik A, Hardiman O. 2009. Screening for replication of genome-wide SNP associations in sporadic ALS. Eur J Hum Genet 17(2):213–8.

Daoud H, Valdmanis PN, Dion PA, Rouleau GA. 2010. Analysis of DPP6 and FGGY as candidate genes for amyotrophic lateral sclerosis. Amyotroph Lateral Scler 11(4):389–91.

DeJesus-Hernandez M, Mackenzie IR, Boeve BF, Boxer AL, Baker M, Rutherford NJ, and others. 2011. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. Neuron 72(2):245–56.

del Aguila MA, Longstreth WT Jr, McGuire V, Koepsell TD, van Belle G. 2003. Prognosis in amyotrophic lateral sclerosis: a population-based study. Neurology 60(5):813–9.

Deng HX, Chen W, Hong ST, Boycott KM, Gorrie GH, Siddique N, and others. 2011. Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. Nature 477(7363):211–5.

Diekstra FP, Saris CGJ, van Rheenen W, Franke L, Jansen RC, van Es MA, and others. 2012. Mapping of gene expression reveals CYP27A1 as a susceptibility gene for sporadic ALS. PLoS One 7(4):e35333.

Diekstra FP, van Vught PWJ, van Rheenen W, Koppers M, Pasterkamp RJ, van Es MA, and others. 2012. UNC13A is a modifier of survival in amyotrophic lateral sclerosis. Neurobiol Aging 33(3):630.e3-e8.

Dunckley T, Huentelman MJ, Craig DW, Pearson JV, Szelinger S, Joshipura K, and others. 2007. Whole-genome analysis of sporadic amyotrophic lateral sclerosis. N Engl J Med 357(8):775–88.

Elden AC, Kim HJ, Hart Mp, Chen-Plotkin AS, Johnson BS, Fang X, and others. 2010. Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. Nature 466(7310):1069–75.

Fallis BA, Hardiman O. 2009. Aggregation of neurodegenerative disease in ALS kindreds. Amyotroph Lateral Scler 10(2):95–8.

Fecto F, Yan J, Vemula SP, Liu E, Yang Y, Chen W, and others. 2011. SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. Arch Neurol 68(11):1440–6.

Figueroa-Romero C, Hur J, Bender DE, Delaney CE, Cataldo MD, Smith AL, and others. 2012. Identification of epigenetically altered genes in sporadic amyotrophic lateral sclerosis. PLoS One 7(12):e52672.

Fogh I, D'Alfonso S, Gellera C, Ratti A, Cereda C, Penco S, and others. 2011. No association of DPP6 with amyotrophic lateral sclerosis in an Italian population. Neurobiol Aging 32(5):966–7.

Fogh I, Ratti A, Gellera C, Lin K, Tiloca C, Moskvina V, and others. 2014. A genome-wide association meta-analysis identifies a novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. Hum Mol Genet 23(8):2220–31.

Forsberg K, Jonsson PA, Andersen PM, Bergemalm D, Graffmo KS, Hultdin M, and others. 2010. Novel antibodies reveal inclusions containing non-native SOD1 in sporadic ALS patients. PLoS One 5(7):e11552.

Gamez J, Caponnetto C, Ferrera L, Syriani E, Marini V, Morales M, and others. 2011. I112M SOD1 mutation causes ALS with rapid progression and reduced penetrance in four Mediterranean families. Amyotroph Lateral Scler 12(1):70–5.

Gamez J, Cervera C, Codina A. 1999. Flail arm syndrome of Vulpian-Bernhart's form of amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 67(2):258.

Garber K. 2008. Genetics. The elusive ALS genes. Science 319(5859):20.

Gijselinck I, Engelborghs S, Maes G, Cuijt I, Peeters K, Mattheijssens M, and others. 2010. Identification of 2 Loci at chromosomes 9 and 14 in a multiplex family with frontotemporal lobar degeneration and amyotrophic lateral sclerosis. Arch Neurol 67(5):606–16.

Gilissen C, Hehir-Kwa JY, Thung DT, van de Vorst M, van Bon BW, Willemsen MH, and others. 2014. Genome sequencing identifies major causes of severe intellectual disability. Nature 511(7509):344–7.

Greenway MJ, Andersen PM, Russ C, Ennis S, Cashman S, Donaghy C, and others. 2006. ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. Nat Genet 38(4):411–3.

Gros-Louis F, Lariviere R, Gowing G, Laurent S, Camu W, Bouchard JP, and others. 2004. A frameshift deletion in

peripherin gene associated with amyotrophic lateral sclerosis. J Biol Chem 279(44):45951–6.

Hadano S, Hand CK, Osuga H, Yanagisawa Y, Otomo A, Devon RS, and others. 2001. A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. Nat Genet 29(2):166–73.

Hardiman O, Greenway M. 2007. The complex genetics of amyotrophic lateral sclerosis. Lancet Neurol 6(4):291–2.

Herdewyn S, Zhao H, Moisse M, Race V, Matthijs G, Reumers J, and others. 2012. Whole-genome sequencing reveals a coding non-pathogenic variant tagging a non-coding pathogenic hexanucleotide repeat expansion in C9orf72 as cause of amyotrophic lateral sclerosis. Hum Mol Genet 21(11):2412–9.

Iida A, Takahashi A, Deng M, Zhang Y, Wang J, Atsuta N, and others. 2011a. Replication analysis of SNPs on 9p21.2 and 19p13.3 with amyotrophic lateral sclerosis in East Asians. Neurobiol Aging 32(4):757.e13–e14.

Iida A, Takahashi A, Kubo M, Saito S, Hosono N, Ohnishi Y, and others. 2011b. A functional variant in ZNF512B is associated with susceptibility to amyotrophic lateral sclerosis in Japanese. Hum Mol Genet 20(18):3684–92.

Johnson JO, Mandrioli J, Benatar M, Abramzon Y, Van Deerlin VM, Trojanowski JQ, and others. 2010. Exome sequencing reveals VCP mutations as a cause of familial ALS. Neuron 68(5):857–64.

Jostins L, Ripke S, Weersma RK, Duerr RH, McGovern DP, Hui KY, and others. 2012. Host-microbe interactions have shaped the genetic architecture of inflammatory bowel disease. Nature 491(7422):119–24.

Kim HJ, Kim NC, Wang YD, Scarborough EA, Moore J, Diaz Z, and others. 2013. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. Nature 495(7442):467–73.

Kwee LC, Liu Y, Haynes C, Gibson JR, Stone A, Schichman SA, and others. 2012. A high-density genome-wide association screen of sporadic ALS in US veterans. PLoS One 7(3):e32768.

Kwiatkowski TJ Jr, Bosco DA, Leclerc AL, Tamrazian E, Vanderburg CR, Russ C, and others. 2009. Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. Science 323(5918):1205–8.

Laaksovirta H, Peuralinna T, Schymick JC, Scholz SW, Lai SL, Myllykangas L, and others. 2010. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. Lancet Neurol 9(10):978–85.

Lagier-Tourenne C, Polymenidou M, Hutt KR, Vu AQ, Baughn M, Huelga SC, and others. 2012. Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. Nat Neurosci 15(11):1488–97.

Landers JE, Melki J, Meininger V, Glass JD, van den Berg LH, van Es MA, and others. 2009. Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. Proc Natl Acad Sci U S A 106(22):9004–9.

Leung CL, He CZ, Kaufmann P, Chin SS, Naini A, Liem RK, and others. 2004. A pathogenic peripherin gene mutation in a patient with amyotrophic lateral sclerosis. Brain Pathol 14(3):290–6.

Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ, and others. 2009. Finding the missing heritability of complex diseases. Nature 461(7265):747–53.

Maruyama H, Morino H, Ito H, Izumi Y, Kato H, Watanabe Y, and others. 2010. Mutations of optineurin in amyotrophic lateral sclerosis. Nature 465(7295):223–6.

Miller RG, Mitchell JD, Moore DH, 2012. Riluzole for amyotrophic lateral sclerosis (ALS)/motor neuron disease (MND). Cochrane Database Syst Rev 3:CD001447.

Mitchell J, Paul P, Chen HJ, Morris A, Payling M, Falchi M, and others. 2010. Familial amyotrophic lateral sclerosis is associated with a mutation in D-amino acid oxidase. Proc Natl Acad Sci U S A 107(16):7556–61.

Mok K, Traynor BJ, Schymick J, Tienari PJ, Laaksovirta H, Peuralinna T, and others. 2012. Chromosome 9 ALS and FTD locus is probably derived from a single founder. Neurobiol Aging 33(1):209.e3–e8.

Momozawa Y, Mni M, Nakamura K, Coppieters W, Almer S, Amininejad L, and others. 2011. Resequencing of positional candidates identifies low frequency IL23R coding variants protecting against inflammatory bowel disease. Nat Genet 43(1):43–7.

Morahan JM, Yu B, Trent RJ, Pamphlett R. 2009. A genome-wide analysis of brain DNA methylation identifies new candidate genes for sporadic amyotrophic lateral sclerosis. Amyotroph Lateral Scler 10(5-6): 418–29.

Morita M, Al-Chalabi A, Andersen PM, Hosler B, Sapp P, Englund E, and others. 2006. A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. Neurology 66(6):839–44.

Munch C, Sedlmeier R, Meyer T, Homberg V, Sperfeld AD, Kurt A, and others. 2004. Point mutations of the p150 subunit of dynactin (DCTN1) gene in ALS. Neurology 63(4):724–6.

Nejentsev S, Walker N, Riches D, Egholm M, Todd JA, 2009. Rare variants of IFIH1, a gene implicated in antiviral responses, protect against type 1 diabetes. Science 324(5925):387–9.

Nishimura AL, Mitne-Neto M, Silva HC, Richieri-Costa A, Middleton S, Cascio D, and others. 2004. A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. Am J Hum Genet 75(5):822–31.

Norris F, Shepherd R, Denys E, Kwei U, Mukai E, Elias L, and others. 1993. Onset, natural history and outcome in idiopathic adult motor neuron disease. J Neurol Sci 118(1):48–55.

O'Roak BJ, Vives L, Girirajan S, Karakoc E, Krumm N, Coe BP, and others. 2012. Sporadic autism exomes reveal a highly interconnected protein network of de novo mutations. Nature 485(7397):246–50.

Oates N, Pamphlett R. 2007. An epigenetic analysis of SOD1 and VEGF in ALS. Amyotroph Lateral Scler 8(2):83–6.

Orlacchio A, Babalini C, Borreca A, Patrono C, Massa R, Basaran S, and others. 2010. SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis. Brain 133(Pt 2):591–8.

Orru S, Manolakos E, Orru N, Kokotas H, Mascia V, Carcassi C, and others. 2012. High frequency of the TARDBP p.Ala382Thr mutation in Sardinian patients with amyotrophic lateral sclerosis. Clin Genet 81(2):172–8.

Orsetti V, Pegoraro E, Cima V, D'Ascenzo C, Palmieri A, Querin G, and others. 2011. Genetic variation in KIFAP3

is associated with an upper motor neuron-predominant phenotype in amyotrophic lateral sclerosis. Neurodegener Dis 8(6):491–5.

Paganoni S, Macklin EA, Lee A, Murphy A, Chang J, Zipf A, and others. 2014. Diagnostic timelines and delays in diagnosing amyotrophic lateral sclerosis (ALS). Amyotroph Lateral Scler Frontotemporal Degener 15(5-6):453–6.

Parkinson N, Ince PG, Smith MO, Highley R, Skibinski G, Andersen PM, and others. 2006. ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). Neurology 67(6):1074–7.

Polymenidou M, Lagier-Tourenne C, Hutt KR, Huelga SC, Moran J, Liang TY, and others. 2011. Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43. Nat Neurosci 14(4):459–68.

Purcell S, Cherny SS, Sham PC. 2003. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. Bioinformatics 19(1):149–50.

Ratti A, Corrado L, Castellotti B, Del Bo R, Fogh I, Cereda C, and others. 2012. C9ORF72 repeat expansion in a large Italian ALS cohort: evidence of a founder effect. Neurobiol Aging 33(10):2528.e7–e14.

Renton AE, Majounie E, Waite A, Simon-Sanchez J, Rollinson S, Gibbs JR, and others. 2011. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. Neuron 72(2):257–68.

Risch N. 1987. Assessing the role of HLA-linked and unlinked determinants of disease. Am J Hum Genet 40(1):1–14.

Risch N. 1990a. Linkage strategies for genetically complex traits. I. Multilocus models. Am J Hum Genet 46(2):222–8.

Risch N. 1990b. Linkage strategies for genetically complex traits. II. The power of affected relative pairs. Am J Hum Genet 46(2): 229–41.

Rivas MA, Beaudoin M, Gardet A, Stevens C, Sharma Y, Zhang CK, and others. 2011. Deep resequencing of GWAS loci identifies independent rare variants associated with inflammatory bowel disease. Nat Genet 43(11):1066–73.

Robberecht W, Aguirre T, Van den Bosch L, Tilkin P, Cassiman JJ, and Matthijs G. 1996. D90A heterozygosity in the SOD1 gene is associated with familial and apparently sporadic amyotrophic lateral sclerosis. Neurology 47(5):1336–9.

Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, and others. 1993. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. Nature 362(6415):59–62.

Shatunov A, Mok K, Newhouse S, Weale ME, Smith B, Vance C, and others. 2010. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. Lancet Neurol 9(10):986–94.

Siddique T, Ajroud-Driss S. 2011. Familial amyotrophic lateral sclerosis, a historical perspective. Acta Myol 30(2):117–20.

Sreedharan J, Blair IP, Tripathi VB, Hu X, Vance C, Rogelj B, and others. 2008. TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. Science 319(5870): 1668–72.

Ticozzi N, Vance C, Leclerc AL, Keagle P, Glass JD, McKenna-Yasek D, and others. 2011. Mutational analysis reveals the FUS homolog TAF15 as a candidate gene for familial amyotrophic lateral sclerosis. Am J Med Genet B Neuropsychiatr Genet 156B(3):285–90.

van Es MA, Dahlberg C, Birve A, Veldink JH, van den Berg LH, Andersen PM. 2010. Large-scale SOD1 mutation screening provides evidence for genetic heterogeneity in amyotrophic lateral sclerosis. J Neurol Neurosurg Psychiatry 81(5):562–6.

van Es MA, Van Vught PW, Blauw HM, Franke L, Saris CG, Andersen PM, and others. 2007. ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. Lancet Neurol 6(10):869–77.

van Es MA, van Vught PWJ, Blauw HM, Franke L, Saris CGJ, Van den Bosch L, and others. 2008. Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis. Nat Genet 40(1):29–31.

van Es MA, Van Vught PWJ, Veldink JH, Andersen PM, Birve A, Lemmens R, and others. 2009a. Analysis of FGGY as a risk factor for sporadic amyotrophic lateral sclerosis. Amyotroph Lateral Scler 10(5-6):441–7.

van Es MA, Veldink JH, Saris CG, Blauw HM, van Vught PW, Birve A, and others. 2009b. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. Nat Genet 41(10):1083–7.

Vance C, Al-Chalabi A, Ruddy D, Smith BN, Hu X, Sreedharan J, and others. 2006. Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3. Brain 129(Pt 4):868–76.

Vance C, Rogelj B, Hortobagyi T, De Vos KJ, Nishimura AL, Sreedharan J, and others. 2009. Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. Science 323(5918):1208–11.

Veldink JH, Kalmijn S, Van der Hout AH, Lemmink HH, Groeneveld GJ, Lummen C, and others. 2005. SMN genotypes producing less SMN protein increase susceptibility to and severity of sporadic ALS. Neurology 65(6):820–5.

Wellcome Trust Case Control Consortium, Craddock N, Hurles ME, Cardin N, Pearson RD, Plagnol V, and others. 2010. Genome-wide association study of CNVs in 16,000 cases of eight common diseases and 3,000 shared controls. Nature 464(7289):713–20.

Wijesekera LC, Mathers S, Talman P, Galtrey C, Parkinson MH, Ganesalingam J, and others. 2009. Natural history and clinical features of the flail arm and flail leg ALS variants. Neurology 72(12):1087–94.

Wu CH, Fallini C, Ticozzi N, Keagle PJ, Sapp PC, Piotrowska K, and others. 2012. Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. Nature 488(7412):499–503.

Xi Z, Zinman L, Moreno D, Schymick J, Liang Y, Sato C, and others. 2013. Hypermethylation of the CpG island near the G4C2 repeat in ALS with a C9orf72 expansion. Am J Hum Genet 92(6):981–9.

Yang Y, Gozen O, Vidensky S, Robinson MB, Rothstein JD. 2010. Epigenetic regulation of neuron-dependent induction of astroglial synaptic protein GLT1. GLIA 58(3):277–86.

Yang Y, Hentati A, Deng HX, Dabbagh O, Sasaki T, Hirano M, and others. 2001. The gene encoding alsin, a protein with three guanine-nucleotide exchange factor domains, is mutated in a form of recessive amyotrophic lateral sclerosis. Nat Genet 29(2):160–5.

Zinman L, Liu HN, Sato C, Wakutani Y, Marvelle AF, Moreno D, and others. 2009. A mechanism for low penetrance in an ALS family with a novel SOD1 deletion. Neurology 72(13):1153–9.