nature neuroscience

State of play in amyotrophic lateral sclerosis genetics

Alan E Renton[1], Adriano Chiò[2] & Bryan J Traynor[1,3]

Considerable progress has been made in unraveling the genetic etiology of amyotrophic lateral sclerosis (ALS), the most common form of adult-onset motor neuron disease and the third most common neurodegenerative disease overall. Here we review genes implicated in the pathogenesis of motor neuron degeneration and how this new information is changing the way we think about this fatal disorder. Specifically, we summarize current literature of the major genes underlying ALS, *SOD1*, *TARDBP*, *FUS*, *OPTN*, *VCP*, *UBQLN2*, *C9ORF72* and *PFN1*, and evaluate the information being gleaned from genome-wide association studies. We also outline emerging themes in ALS research, such as next-generation sequencing approaches to identify *de novo* mutations, the genetic convergence of familial and sporadic ALS, the proposed oligogenic basis for the disease, and how each new genetic discovery is broadening the phenotype associated with the clinical entity we know as ALS.

ALS is an adult-onset neurodegenerative disorder characterized by rapidly progressive paralysis and death from respiratory failure, typically within 2 to 3 years of symptom onset[1]. There are no effective cures for ALS, though the benzothiazole riluzole slows the rate of progression and prolongs survival by three months[1]. As in other neurodegenerative diseases, ~10% of ALS is classified as familial, whereas the remaining 90% of cases are considered sporadic, as they appear to occur randomly throughout the community[1].

Unraveling the genetic etiology of ALS has provided fundamental insights into the cellular mechanisms underlying neuron degeneration, as well as facilitating disease modeling and the design and testing of targeted therapeutics; thus, it is not surprising that considerable resources have been devoted to finding pathogenic mutations. *SOD1* mutations were found to cause familial ALS in 1993 (ref. 2), but there was a long hiatus until the discovery of the next ALS gene, *TARDBP*, in 2008 (ref. 3). Today, the pace of gene discovery has greatly accelerated, fuelled in large part by advances in sequencing and genotyping technology. In the last 3 years alone, six new ALS genes have been discovered[4–10], and we now know the genetic etiology of two-thirds of familial cases and about 11% of sporadic ALS cases.

This Review will focus on recent genetic discoveries in ALS and how genetics is changing our understanding of this fatal, mysterious disorder. In the first section, we examine the main genes implicated in the pathogenesis of ALS. We will then discuss attempts to find genetic variants underlying sporadic ALS using genome-wide association studies (GWAS). Finally, we will describe emerging concepts in ALS genetics and where we ultimately think the road ahead leads.

[1]Neuromuscular Diseases Research Unit, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, Maryland, USA. [2]Rita Levi Montalcini Department of Neuroscience, University of Turin, Turin, Italy. [3]Department of Neurology, Brain Sciences Institute, Johns Hopkins University, Baltimore, Maryland, USA. Correspondence should be addressed to B.J.T. (traynorb@mail.nih.gov).

Received 23 June; accepted 22 October; published online 26 December 2013; doi:10.1038/nn.3584

Familial ALS genes

To provide a historical context, we present the main ALS genes in the order in which they were discovered (**Fig. 1** and **Table 1**).

Superoxide dismutase 1 (*SOD1*). The identification of dominant missense mutations in *SOD1* 20 years ago represented the first demonstration that linkage analysis could successfully pinpoint the underlying genetic cause of a rare neurodegenerative disease[2]. Although over a hundred *SOD1* mutations have been reported, reliable genetic evidence of pathogenicity exists for only a portion of these[11]. Nevertheless, population-based studies show that mutations in this gene account for ~12% of familial cases and ~1% of sporadic cases[12].

Considerable phenotypic heterogeneity occurs across the various *SOD1* mutations. For example, the A4V mutation, which is the most frequent variant in North America, gives rise to an aggressive form of ALS that typically leads to death within a year after symptom onset[13]. In contrast, the homozygous D90A mutation in the same gene is associated with an indolent course, with patients developing respiratory failure only after 10 years of illness[14]. Cognitive impairment is not a prominent feature of *SOD1* disease, though patients with D90A manifest cognitive issues at the later stages of the disease, perhaps reflecting their protracted survival and the corresponding longer time to manifest disease spreading to nonmotor prefrontal areas[15].

The discovery that SOD1 mutations cause ALS led directly to the development of the *SOD1* transgenic mouse. Though important in elucidating the cellular mechanisms by which disruption of this gene predispose to motor neuron degeneration[16], the use of this model to select agents for human trials has been increasingly called into question[17]. Indeed, the pathology of human *SOD1* ALS is now thought to be distinct from that of all other types of ALS, in that it lacks the TDP-43 and/or FUS pathology present in nearly every other instance[18]. In recognition of this fact, there have been recent efforts to selectively treat patients carrying mutated *SOD1* using antisense oligonucleotide therapy designed to knock down expression of the gene[19].

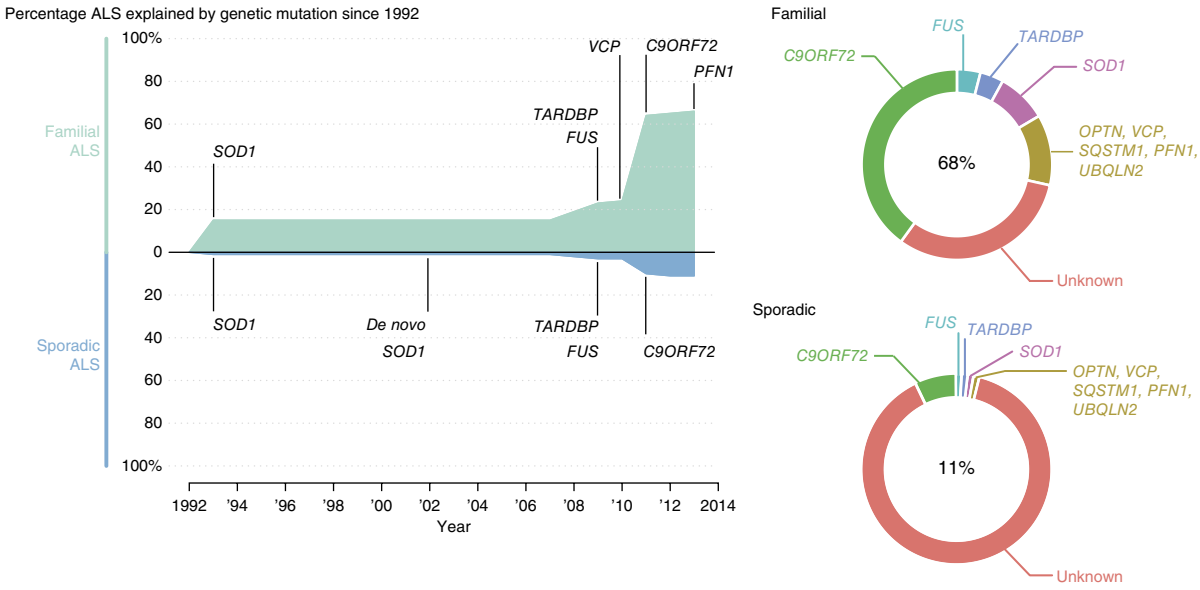

Figure 1 Timeline of gene discoveries in familial and sporadic ALS. Values represent the proportion of ALS explained by each gene in populations of European ancestry. References are provided in the main text.

TAR DNA-binding protein (*TARDBP*). A landmark event in our understanding of ALS pathogenesis was the discovery that the TDP-43 protein is a major component of the ubiquitin-positive neuronal inclusions that are the pathological hallmark of both ALS and frontotemporal dementia (FTD)[20]. This seminal observation provided vital evidence that these distinct conditions constitute a disease spectrum, rather than being discrete clinical entities. It also directly led to the discovery of mutations in *TARDBP* (which encodes TDP-43) in autosomal dominant ALS and FTD families[3,21]. These pathogenic variants were mostly located in the C terminus, which is involved in ribonucleoprotein binding and splicing.

Mutations in *TARDBP* account for ~4% of familial ALS cases and a smaller percentage of sporadic ALS cases[22]. Although *TARDBP* mutations are found in ALS families across the globe, some regional variability does exist. For example, the A382T mutation of the protein is particularly frequent in Sardinia, reflecting the conserved nature of that island population combined with a historical founder effect[23]. Nevertheless, the overall mutational frequency of the gene remains much lower than the occurrence of TDP-43 neuropathological changes observed in autopsies.

The discovery of the central role of TDP-43 in ALS pathogenesis has highlighted the importance of RNA processing. Additional support for this hypothesis comes from the discovery of *FUS*, another RNA-binding protein (see below). Considerable efforts have been made to understand exactly how mutations in these genes disrupt RNA transcription and to identify RNA species modified by both proteins[24].

Fused in sarcoma (*FUS*). Shortly after the cloning of *TARDBP*, missense mutations of *FUS* were identified as the cause of chromosome 16p–linked familial ALS[25,26]. Although this gene accounts for a small percentage of familial cases (~4%), its discovery caused considerable excitement in the field because the protein shares functional homology with TDP-43. Furthermore, *FUS* mutations cluster in the RNA-binding domain at the C terminus of the protein, as seen with *TARDBP*. These observations reinforced the importance of abnormal RNA metabolism in motor neuron degeneration.

Neuropathological analyses of patients carrying *FUS* mutations have yielded important insights into the disease. Though some debate remains, it would appear that *FUS* cases are characterized by FUS-immunoreactive cytoplasmic inclusions with a striking absence of the ubiquitin-positive and TDP-43-positive aggregates that distinguish most ALS cases[25,26]. One possible explanation for this finding is that FUS acts downstream of TDP-43 in the same pathway[27].

Optineurin (*OPTN*). *OPTN* mutations were initially described as a cause of autosomal recessive ALS in Japanese families in 2010 (ref. 9). Since then, it has become clear that heterozygous mutations in this gene rarely cause familial ALS among people of European ancestry[22]. This may reflect different genetic etiologies underlying motor neuron degeneration across major ethnic groups. Indeed, autosomal recessive ALS may be more common in Japan compared to Europe and North America.

Nevertheless, the phenotypic pleiotropy associated with *OPTN* is intriguing. Mutations in this gene have long been described as a cause

Table 1 Genes known to carry ALS-causing mutations

Gene	Location	Inheritance	Percentage explained		Putative protein function
			Familial ALS	Sporadic ALS	
TARDBP	1p36	AD	4	1	RNA metabolism
SQSTM1	5q35	AD	1	<1	Ubiquitination; autophagy
C9ORF72	9p21	AD	40	7	DENN protein
VCP	9p13	AD	1	1	Proteasome; vesicle trafficking
OPTN	10p13	AR and AD	<1	<1	Vesicle trafficking
FUS	16p11	AD and AR	4	1	RNA metabolism
PFN1	17p13	AD	<1	<1	Cytoskeletal dynamics
SOD1	21q22	AD and AR	12	1–2	Superoxide metabolism
UBQLN2	Xp11	XD	<1	<1	Proteasome

Values represent the percentage of ALS explained by each gene in populations of European ancestry. References are provided in the main text. AD, autosomal dominant; AR, autosomal recessive; XD, X-linked dominant; DENN, differentially expressed in normal and neoplasia.

of primary open angle glaucoma[28], and more recently the *OPTN* locus was implicated in a GWAS of Paget's disease of bone[29]. Along with *VCP* and *SQSTM1* (see below), this provides information supporting a clinical overlap between Paget's disease and ALS. Although *OPTN* regulates diverse cellular processes, including membrane trafficking, protein secretion, cell division and host defense against pathogens[30], it remains to be determined exactly how mutations in this gene may give rise to such a wide range of clinical phenotypes.

Valosin-containing protein (*VCP*). Also in 2010, we reported that mutations in *VCP* are responsible for 1–2% of familial ALS cases[5], a finding that was subsequently confirmed by other studies[31,32]. This was the first time that whole-exome sequencing had been successfully applied to a major neurodegenerative disorder. The power of this new technology is based on its use of the massively parallel sequencing capabilities of next-generation sequencing platforms to rapidly identify rare variants in the ~2% of the genome that encodes proteins. Moreover, whole-exome sequencing is a realistic strategy for detecting pathogenic variants in small families where linkage analysis would not be possible due to a shortage of DNA samples from affected individuals. This is particularly true in ALS, where the rapidly fatal nature of the syndrome makes it challenging to collect families of sufficient size to perform linkage analysis.

Mutations in *VCP* were already known to underlie an unusual clinical syndrome characterized by FTD, inclusion body myopathy and Paget's disease of the bone (IBMPFD)[33]. Thus, the discovery of mutations in the same gene resulting in an ALS phenotype represented another step in our unraveling of the genetic links between motor neuron degeneration and FTD. The coexistence of inclusion body myopathy in these patients was similarly interesting because it demonstrated that mutations in a single gene could result in pathology on both sides of the neuromuscular junction. This has given rise to the concept of multisystem proteinopathy, in which multiple tissues are affected[34]. Recent work has convincingly demonstrated that mutations in *VCP* cause mitochondrial uncoupling leading to a significant reduction of cellular ATP production, perhaps explaining disparate effects across multiple tissues[35].

Ubiquilin 2 (*UBQLN2*). Missense mutations in *UBQLN2*, located on the short arm of chromosome X, were initially identified in apparently autosomal dominant ALS pedigrees lacking male-to-male transmission[4]. Although isolated ALS was the predominant phenotype associated with mutations in this gene, occasional patients had concomitant symptoms of FTD. It is already evident that *UBQLN2* mutations are not a common cause of familial ALS[36]. Nevertheless, ubiquilin 2 pathology has been observed in ALS patients who do not carry mutations in the gene, suggesting that this protein, which regulates proteasome degradation of ubiquitinated proteins, may be an important component of the final common pathway mediating motor neuron degeneration[4].

Hexanucleotide repeat expansion in *C9ORF72*. We and others recently reported that a massive hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21–linked ALS and FTD[7,8]. Long sought after, cloning of this locus has reinvigorated the ALS and FTD research field for a variety of reasons. First, the pathogenic expansion accounts for a remarkable percentage of both familial ALS (~40%) and familial FTD (~25%) and genetically explains the majority of the overlap of these two disparate clinical syndromes[37]. Second, the repeat expansion has been found to account ~7% of apparently sporadic ALS cases in people of European ancestry, marking the first time that a genetic etiology has been identified for more than just the occasional sporadic case. Third, this is the first time that a large intronic repeat expansion has been implicated in ALS. Given that repeat expansions are known to disrupt RNA metabolism in other neurodegenerative diseases, it again points toward the importance of this pathway in the maintenance of motor and frontal cortex neurons. Finally, there is a real chance that gene therapy based on antisense oligonucleotides will be effective in slowing progression in ALS patients who carry the expansion (approximately one in ten ALS cases in European-ancestry populations). There are already concerted efforts to bring such a therapy forward into human trials.

Several important questions concerning *C9ORF72* remain unanswered. For example, what is the full range of phenotypes associated with this mutation? The expansion may underlie a small portion of clinical Alzheimer's disease and Parkinson's disease cases[38,39], but what is its role in neuropsychiatric disorders? A corollary question is why some patients present with motor dysfunction and others with frontal lobe impairment. Although there are anecdotal reports of anticipation (the onset of symptoms at an earlier age in subsequent generations) in *C9ORF72* families[40], is this a robust finding? If so, is the repeat expansion itself unstable, and what role does such instability play in determining disease phenotype and severity? Early data point to considerable instability of the expansion both across families and within individuals (somatic mutation)[41]. Is the GGGGCC repeat pure or interrupted, and how might that influence phenotype? What genomic and cellular substrates underlie the variable penetrance observed among individuals carrying the expansion?[37] Methylation has been shown to be an important determinant of age of symptom onset in other repeat-expansion diseases, and there are data to suggest that this may be relevant in *C9ORF72* (ref. 42).

Finally, and perhaps most pressing for the development of gene therapies targeting the locus, how does the repeat expansion give rise to disease? RNA foci are observed in fibroblasts derived from *C9ORF72* patients, implicating disruption of RNA metabolism through sequestration of RNA-binding proteins and other RNA species[8]. However, a more direct role for C9ORF72 protein in the nucleus has not been excluded, suggesting that haploinsufficiency may also be relevant. *C9ORF72* hexanucleotide repeats form highly stable RNA G-quadruplexes, which may influence telomere stability and RNA transcription, splicing, translation and transport[43]. More recently, repeat-associated non-ATG (RAN) protein translation has been observed[44]. Although it is not known whether the inclusions generated from such aberrant translation are involved in neuronal cell death, this process has been directly implicated in cellular toxicity associated with fragile X tremor ataxia, a different noncoding repeat expansion disease[45].

Sequestosome 1 (*SQSTM1*). *SQSTM1* encodes p62, a major pathologic protein deposited in neurodegeneration. P62 regulates ubiquitin binding and activation of nuclear factor kappa-B signaling. Mutations in this gene are known to cause Paget's disease of bone[46]. A candidate gene screening approach identified missense and deletion variants in ~1% of ALS cases[10]. Follow-up studies have shown a similar mutational frequency rate, though, as of yet, segregation of variants with disease within a large family has not been reported[47,48].

Profilin 1 (*PFN1*). In late 2012 Wu *et al.* reported that missense mutations in *PFN1* segregated with disease in several large ALS kindreds[6]. Although this gene has only been recently described as pathogenic mutational screening of sizeable ALS and FTD cohorts from several populations already demonstrate that *PFN1* mutations are not

a prominent cause of neurodegeneration[49]. Nevertheless, this discovery implicates a new cellular mechanism in the pathogenesis of ALS, namely disruption of the cytoskeletal architecture of the neuron.

Other genes. Mutations in several other genes have been reported as rare causes of ALS or ALS-like syndromes (**Table 2**). In the case of *ALS2*, the associated phenotype is more consistent with slowly progressive, juvenile-onset, hereditary spastic paraplegia, rather than adult-onset ALS[50,51]. Similarly, mutations in *SETX* are more typically associated with an ataxia phenotype, rather than ALS[52,53].

Table 2 Other genes implicated in the pathogenesis of ALS

Gene	Location	Inheritance	Predominant clinical syndromes	Putative protein function
DCTN1	2p13	AD	PMA; Perry syndrome	Axonal transport
ALS2	2q33	AR	Juvenile PLS; infantile HSP	Vesicle trafficking
CHMP2B	3p11	AD	Familial ALS; sporadic ALS; FTD	Vesicle trafficking
FIG4	6q21	AD and AR	CMT; familial ALS	Vesicle trafficking
HNRNPA2B1	7p15	AD	Multisystem proteinopathy; ALS	RNA metabolism
ELP3	8p21	Undefined	Sporadic ALS	RNA metabolism
SETX	9q34	AD	Juvenile ALS; ataxia with oculomotor apraxia	RNA metabolism
HNRNPA1	12q13	AD	Multisystem proteinopathy; ALS	RNA metabolism
ATXN2	12q24	Undefined	Sporadic ALS; ataxia	Endocytosis; RNA translation
ANG	14q11	AD	Familial ALS; sporadic ALS	Angiogenesis
SPG11	15q14	AR	Juvenile ALS; HSP	DNA damage repair
VAPB	20q13	AD	PMA; FALS	Vesicle trafficking
NEFH	22q12	AD	Familial ALS; sporadic ALS	Axonal transport

AD, autosomal dominant; AR, autosomal recessive; CMT, Charcot-Marie-Tooth disease; HSP, hereditary spastic paraplegia; PLS, primary lateral sclerosis; PMA, progressive muscular atrophy.

In other cases, the genetic evidence supporting a role as an ALS gene is not fully convincing. For example, there have been no reports of *FIG4* or *DAO* mutations as a cause of ALS since their initial publications[54,55]. By the same token, only a single pedigree showing convincing segregation of ALS with a *DCTN1* mutation has been described, and mutations in this gene are now more commonly linked to Parkinson's disease with hypoventilation and depression (Perry syndrome)[56,57]. Apart from the Brazilian ALS families in which the original *VAPB* mutations were reported, there have been no additional publications of segregating mutations in this gene[58,59]. Though there is good evidence for a private mutation in *CHMP2B* in a large Danish family with FTD, there is only minimal data to support its involvement in ALS pathogenesis[60–62]. The role of the candidate gene *ANG*, nominated on the basis of its angiogenic properties, remains ambiguous for both familial and sporadic ALS[22,63–65].

Genome-wide association studies of ALS

Susceptibility loci. There have been 14 GWAS published in ALS, and they have already made significant contributions to our understanding of ALS genetics[66–79]. Our own GWAS of ALS in Finland was the first to identify a significant association peak on the short arm of chromosome 9 in this founder population[75]. This remains the most consistent signal yet observed[73,76,78] and was instrumental in the cloning of the *C9ORF72* locus, as the dense nature of the single nucleotide polymorphisms (SNPs) on the genotyping platform narrowed the scope of the search.

A variety of other risk loci have been nominated on the basis of ALS GWAS. Many of them, such as *FGGY*, *ITPR2* and *DPP6* have not stood up to replication in large cohorts[71,77,80]. Others, such as *UNC13A*, appear more robust, although true replication of this locus in an independent cohort is still pending[78,81]. More recently, a GWAS of ALS in people of non-European ancestry identified new loci at 1q32 and 22p11 as potentially relevant to ALS pathogenesis, though the discovery cohort was small (n = 506) for an outbred population such as the Han Chinese[79]. The field eagerly anticipates independent replication of these findings in larger case-control cohorts.

ALS genetic research has been largely focused on populations of European ancestry, and attention has only recently shifted to other ethnicities. Early indications are that the genetic architecture of ALS is distinct in other regions of the world. For example, the frequency of the *C9ORF72* repeat expansion is far lower among Japanese than among Europeans[37], whereas the inverse pattern is observed for *OPTN* mutations. Future genetic discoveries in non-European populations will likely unravel the diverse means by which motor neuron degeneration can occur.

Age at symptom onset and survival phenotypes. In addition to searching for susceptibility loci, there have been genome-wide efforts to identify genetic variants influencing ALS phenotype. For example, genetic variation in the *KIFAP3* and *EPHA4* loci has been reported to significantly influence survival among ALS patients[72,82], and a large meta-analysis of ALS GWAS recently reported that age of onset is modulated by a locus on the short arm of chromosome 1 (ref. 78). The motivation for these efforts is based on the notion that such loci represent final common pathways involved in the death of motor neurons and that the cellular mechanisms that drive neurodegeneration forward may be distinct from the initiating events. Final common pathways may also be more attractive targets for therapeutic intervention, as such agents are more likely to be effective across the gamut of ALS cases. Despite this, caution against overinterpretation of such secondary phenotypes is required. Indeed, attempts to replicate the effect of the *KIFAP3* locus on survival have met with varied success[74].

Lessons learned from ALS GWAS. The current state of play in ALS illustrates a recurring theme with GWAS methodology, namely that the sheer number of association tests performed on the same data set often leads to false-positive findings[83]. Although a variety of statistical tools can be employed to counter this, independent replication remains the gold standard. Such replication studies are greatly facilitated by the public availability of raw genotype data generated in earlier GWAS. The release of such data allows other researchers around the world to access, reanalyze and combine it with their own GWAS, thereby increasing the power of their studies at no added cost.

By their very nature, GWAS of outbred populations require several thousand case and control samples to have sufficient power to identify risk loci. The Coriell ALS DNA Repository (http://www.coriell.org/), funded by the ALS Association, the Muscular Dystrophy Association and the National Institute of Neurological Disorders and Stroke, distributes well-phenotyped biological samples from ~2,000 case and ~6,000 population control subjects[84]. The availability of this resource has invigorated the field by lowering the barriers for laboratories to engaging in genetic research.

The discovery of more and more causative genes confirms that ALS is not a monolithic clinical entity but rather consists of a group of diseases unified by the common theme of progressive motor neuron degeneration. Such phenotypic and genetic heterogeneity confounds the ability of GWAS to identify associated regions. A standard approach to overcoming this obstacle has been to increase the size of case-control cohorts genotyped as part of GWAS. This approach

has been successful in other neurodegenerative diseases, such as Parkinson's disease and Alzheimer's disease[85,86], and lends credence to the view that similarly sized studies are likely to reveal more loci important to the pathogenesis of ALS.

A complementary technique has been to study conserved populations that display elevated rates of ALS. The reduced genetic heterogeneity found in such communities dramatically increases power to identify new genes responsible for disease. In Finland, our GWAS identified association signals on chromosome 9 (corresponding to the pathogenic *C9ORF72* repeat expansion) and on chromosome 21 (corresponding to the D90A allele of *SOD1*)[75]. These loci account for nearly all familial ALS cases in Finland and explain the high disease incidence observed in that population isolate. Our studies of Sardinian ALS patients have revealed two causative genes operating in that island population, namely the A382T-encoding allele of *TARDBP* and the *C9ORF72* repeat expansion[23,37]. Similar efforts to exploit the Irish population have been less successful, but this may be due to the small size of the cohort or because this country does not have a higher rate of ALS than other European populations[69,87].

Future directions for ALS GWAS. Older versions of genome-wide arrays selected common SNPs (minor allele frequency >5% in the general population) to capture the maximum variation across the genome. This allowed the common disease/common variant hypothesis to be tested for a particular disease. In contrast, newer genotyping platforms (for example, the HumanExome Beadchip from Illumina) focus on rare variants (minor allele frequency <5%) located in the coding portion of the genome. GWAS performed with these chips will test the hypothesis that a disease is caused by multiple rare variants. Given the relative rarity of ALS in the general population (incidence ~2 per 100,000 in Europe), these new types of GWAS hold great promise.

Emerging themes in ALS genetics

***De novo* mutations.** Spontaneously occurring mutations are a well-known cause of neurological and non-neurological conditions, such as neurofibromatosis type 1 (ref. 88) and Hirschsprung's disease[89]. Indeed, *de novo* mutations of *FUS* and *SOD1* have been described in sporadic ALS cases[90–92]. Exome sequencing of parent–case offspring trios offers a new method to systematically search the genome for such variants. A recent publication from Gitler and colleagues highlights the potential power of this method and led to the identification and functional characterization of mutations in the chromatin regulator *SS18L1* (also known as *CREST*) as possibly underlying ALS[93]. This publication represents an excellent first step, though it should be recognized that not all *de novo* variants are necessarily pathogenic. Validation of these 'hits' in larger cohorts of familial and sporadic ALS cases is necessary to ensure that this new technique really flourishes.

Breakdown in classification of familial and sporadic ALS. The discovery that *C9ORF72* is responsible for a sizeable proportion of sporadic ALS cases in populations of European ancestry has led to a fundamental shift in our understanding of the disease. Although categorization of ALS cases as familial or sporadic on the basis of family history retains some utility, it is clear that the underlying biology is more nuanced and that such clinical classification should not be viewed in absolute terms[94]. This has profound implications for the clinical care and genetic testing in ALS[95]. It also lends credence to the notion that a genetic component underlies all of ALS and suggests that environmental and stochastic factors act as precipitating factors in genetically susceptible individuals, rather than being directly causative in their own right.

Oligogenic basis for ALS. An oligogenic basis has recently been proposed in ALS families exhibiting Mendelian inheritance[96]. Under this model, mutations in two or more genes are required for an individual to develop disease. In some ways, this is reminiscent of the Knudson two-hit hypothesis implicated in carcinogenesis. While such events can occur in isolated populations, these are likely to be chance events driven by the high frequency of founder mutations in those regions[97]. A more parsimonious explanation for finding two 'mutations' in an ALS patient from an outbred population is that one of these variants is not truly pathogenic. In that regard, it is telling that many of the patients reported to carry two mutations in the Dutch population consisted of a mutation in an ALS gene and what may be a benign polymorphism in *ANG*[96].

Nevertheless, one cannot a priori discount the possibility that ALS patients carry a primary genetic variant that drives susceptibility to disease and a secondary locus that influences age of symptom onset, symptom presentation and/or disease progression. Genome-wide data from several thousand patients across multiple countries will be required to resolve whether such genetic oligogenicity is relevant to ALS. The public availability of such large-scale data sets will allow future researchers to test for epistatic interactions among genes.

Extension of ALS phenotype and overlap between ALS and FTD. The more we learn about the genetics of ALS, the broader the associated phenotype becomes. For example, recognizing that ALS and FTD represent overlapping clinical syndromes has been a major step forward in our understanding, and this convergence has been strengthened by the discovery of *C9ORF72* and other genes. In addition, *VCP* and *SQSTM1* show that the condition, at least in some cases, is not limited to the CNS and that diverse tissues such as muscle and bone may be affected. The discovery of mutations in *HNRNPA2B1* and *HNRNPA1* in families with ALS and in families with muscle-, bone- and neurodegeneration reinforces the notion that at least some forms of ALS are part of a more widespread disease process[34]. The term "multisystem proteinopathy" has been suggested to describe these diverse findings[34].

Genes involved in other forms of motor neuron disease, such as hereditary spastic paraplegia and Charcot-Marie-Tooth disease, and even forms of neurodegeneration that are now regarded as quite distinct, may ultimately be linked to the ALS phenotype rubric. Expansions of the polyglutamine repeat in the coding region of *ATXN2* cause a form of spinocerebellar ataxia (SCA2), and there appears to be clinical overlap between ALS and SCA2, with intermediate-length expansions reported to increase the risk of ALS[98,99]. Intriguingly, *ANG* has been implicated in the pathogenesis of Parkinson's disease[100], although it is unclear why this locus was not detected in a previous GWAS of Parkinson's disease involving 12,000 cases and 20,000 controls[86].

Conclusions and future directions

We have come a long way since the discovery of the first ALS gene 20 years ago, and there is no doubt that genomics will continue to drive the research field forward. Analysis of increasingly large genetic data sets will improve our understanding of neurodegeneration. Although exome sequencing is in fashion, whole-genome sequencing will rise to prominence as costs continue to drop. Genome sequencing will yield even greater insight into the genetic architecture of ALS by providing a complete catalog of rare variants present in cases and allowing us to explore the role of noncoding and intergenic genetic variation in the pathogenesis of motor neuron degeneration. Combined with the promise of effective gene therapies, it is plausible that the next 20 years will see the mysteries of ALS simply melt away.

ACKNOWLEDGMENTS

This work was supported in part by the Intramural Research Programs of the US National Institutes of Health, National Institute on Aging (Z01-AG000949-02) and National Institute of Neurological Disorders and Stroke. The work was also supported by the Packard Center for ALS Research at Hopkins (B.J.T.), the ALS Association (B.J.T., A.C.), Microsoft Research (B.J.T.), AriSLA (B.J.T., A.C.), the Italian Health Ministry (Ricerca Sanitaria Finalizzata 2007 to A.C.), Fondazione Vialli e Mauro ONLUS (A.C.), Federazione Italiana Giuoco Calcio (A.C., B.J.T.), Compagnia di San Paolo (A.C.) and the European Community's Health Seventh Framework Programme under grant agreement 259867 (A.C.).

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details are available in the online version of the paper.

1. Rowland, L.P. & Shneider, N.A. Amyotrophic lateral sclerosis. *N. Engl. J. Med.* **344**, 1688–1700 (2001).
2. Rosen, D.R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59–62 (1993).
This study was the first to identify a genetic cause of familial ALS.
3. Sreedharan, J. *et al.* TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science* **319**, 1668–1672 (2008).
This study identified mutations in *TARDBP*, which encodes the TDP-43 protein, as a cause of familial ALS.
4. Deng, H.X. *et al.* Mutations in *UBQLN2* cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477**, 211–215 (2011).
5. Johnson, J.O. *et al.* Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* **68**, 857–864 (2010).
This study was the first to apply exome sequencing to identify a genetic cause of familial ALS and represents an initial step in unraveling the genetic overlap between ALS and FTD.
6. Wu, C.H. *et al.* Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature* **488**, 499–503 (2012).
7. Renton, A.E. *et al.* A hexanucleotide repeat expansion in *C9ORF72* is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257–268 (2011).
8. DeJesus-Hernandez, M. *et al.* Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron* **72**, 245–256 (2011).
Refs. 7 and 8 identified the hexanucleotide repeat expansion in *C9ORF72* as a major cause of familial and sporadic ALS and FTD.
9. Maruyama, H. *et al.* Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* **465**, 223–226 (2010).
10. Fecto, F. *et al.* SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Arch. Neurol.* **68**, 1440–1446 (2011).
11. Andersen, P.M. Amyotrophic lateral sclerosis associated with mutations in the CuZn superoxide dismutase gene. *Curr. Neurol. Neurosci. Rep.* **6**, 37–46 (2006).
12. Chiò, A. *et al.* Prevalence of SOD1 mutations in the Italian ALS population. *Neurology* **70**, 533–537 (2008).
13. Cudkowicz, M.E. *et al.* Epidemiology of mutations in superoxide dismutase in amyotrophic lateral sclerosis. *Ann. Neurol.* **41**, 210–221 (1997).
14. Andersen, P.M. *et al.* Autosomal recessive adult-onset amyotrophic lateral sclerosis associated with homozygosity for Asp90Ala CuZn-superoxide dismutase mutation. A clinical and genealogical study of 36 patients. *Brain* **119**, 1153–1172 (1996).
15. Turner, M.R. *et al.* Distinct cerebral lesions in sporadic and 'D90A' SOD1 ALS: studies with [11C]flumazenil PET. *Brain* **128**, 1323–1329 (2005).
16. Pasinelli, P. *et al.* Amyotrophic lateral sclerosis-associated SOD1 mutant proteins bind and aggregate with Bcl-2 in spinal cord mitochondria. *Neuron* **43**, 19–30 (2004).
17. Ludolph, A.C. *et al.* Guidelines for the preclinical in vivo evaluation of pharmacological active drugs for ALS/MND: report on the 142nd ENMC international workshop. *Amyotroph. Lateral Scler.* **8**, 217–223 (2007).
18. Mackenzie, I.R. *et al.* Pathological TDP-43 distinguishes sporadic amyotrophic lateral sclerosis from amyotrophic lateral sclerosis with SOD1 mutations. *Ann. Neurol.* **61**, 427–434 (2007).
19. Miller, T.M. *et al.* An antisense oligonucleotide against SOD1 delivered intrathecally for patients with SOD1 familial amyotrophic lateral sclerosis: a phase 1, randomised, first-in-man study. *Lancet Neurol.* **12**, 435–442 (2013).
This study was the first report of gene therapy using antisense oligonucleotides in ALS patients.
20. Neumann, M. *et al.* Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Science* **314**, 130–133 (2006).
This seminal paper identified TDP-43 as a major component of the ubiquitin-positive neuronal inclusions that are the neuropathological hallmark of ALS and FTD.
21. Chiò, A. *et al.* Amyotrophic lateral sclerosis-frontotemporal lobar dementia in 3 families with p.Ala382Thr TARDBP mutations. *Arch. Neurol.* **67**, 1002–1009 (2010).
22. Chiò, A. *et al.* Extensive genetics of ALS: a population-based study in Italy. *Neurology* **79**, 1983–1989 (2012).
23. Chiò, A. *et al.* Large proportion of amyotrophic lateral sclerosis cases in Sardinia due to a single founder mutation of the TARDBP gene. *Arch. Neurol.* **68**, 594–598 (2011).
24. Lagier-Tourenne, C. *et al.* Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs. *Nat. Neurosci.* **15**, 1488–1497 (2012).
25. Kwiatkowski, T.J. Jr. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science* **323**, 1205–1208 (2009).
26. Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science* **323**, 1208–1211 (2009).
27. Kabashi, E. *et al.* FUS and TARDBP but not SOD1 interact in genetic models of amyotrophic lateral sclerosis. *PLoS Genet.* **7**, e1002214 (2011).
28. Rezaie, T. *et al.* Adult-onset primary open-angle glaucoma caused by mutations in optineurin. *Science* **295**, 1077–1079 (2002).
29. Albagha, O.M. *et al.* Genome-wide association study identifies variants at *CSF1*, *OPTN* and *TNFRSF11A* as genetic risk factors for Paget's disease of bone. *Nat. Genet.* **42**, 520–524 (2010).
30. Kachaner, D., Génin, P., Laplantine, E. & Weil, R. Toward an integrative view of Optineurin functions. *Cell Cycle* **11**, 2808–2818 (2012).
31. Koppers, M. *et al.* VCP mutations in familial and sporadic amyotrophic lateral sclerosis. *Neurobiol. Aging* **33**, 837.e7–837.e13 (2012).
32. DeJesus-Hernandez, M. *et al.* Novel p.Ile151Val mutation in VCP in a patient of African American descent with sporadic ALS. *Neurology* **77**, 1102–1103 (2011).
33. Watts, G.D. *et al.* Inclusion body myopathy associated with Paget disease of bone and frontotemporal dementia is caused by mutant valosin-containing protein. *Nat. Genet.* **36**, 377–381 (2004).
34. Kim, H.J. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467–473 (2013).
35. Bartolome, F. *et al.* Pathogenic VCP mutations induce mitochondrial uncoupling and reduced ATP levels. *Neuron* **78**, 57–64 (2013).
36. Daoud, H. *et al.* UBQLN2 mutations are rare in French and French-Canadian amyotrophic lateral sclerosis. *Neurobiol. Aging* **33**, 2230 e1–2230.e5 (2012).
37. Majounie, E. *et al.* Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *Lancet Neurol.* **11**, 323–330 (2012).
38. Lesage, S. *et al.* C9orf72 repeat expansions are a rare genetic cause of parkinsonism. *Brain* **136**, 385–391 (2013).
39. Majounie, E. *et al.* Repeat expansion in C9ORF72 in Alzheimer's disease. *N. Engl. J. Med.* **366**, 283–284 (2012).
40. Chiò, A. *et al.* Clinical characteristics of patients with familial amyotrophic lateral sclerosis carrying the pathogenic GGGGCC hexanucleotide repeat expansion of C9ORF72. *Brain* **135**, 784–793 (2012).
41. Beck, J. *et al.* 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 (2013).
42. Xi, Z. *et al.* Hypermethylation of the CpG island near the G_4C_2 repeat in ALS with a C9orf72 expansion. *Am. J. Hum. Genet.* **92**, 981–989 (2013).
43. Fratta, P. *et al.* C9orf72 hexanucleotide repeat associated with amyotrophic lateral sclerosis and frontotemporal dementia forms RNA G-quadruplexes. *Sci. Rep.* **2**, 1016 (2012).
44. Mori, K. *et al.* The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTLD/ALS. *Science* **339**, 1335–1338 (2013).
45. Todd, P.K. *et al.* CGG repeat-associated translation mediates neurodegeneration in fragile X tremor ataxia syndrome. *Neuron* **78**, 440–455 (2013).
46. Laurin, N., Brown, J.P., Morissette, J. & Raymond, V. Recurrent mutation of the gene encoding sequestosome 1 (*SQSTM1*/*p62*) in Paget disease of bone. *Am. J. Hum. Genet.* **70**, 1582–1588 (2002).
47. Teyssou, E. *et al.* Mutations in SQSTM1 encoding p62 in amyotrophic lateral sclerosis: genetics and neuropathology. *Acta Neuropathol.* **125**, 511–522 (2013).
48. Rubino, E. *et al.* SQSTM1 mutations in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Neurology* **79**, 1556–1562 (2012).
49. Lattante, S., Le Ber, I., Camuzat, A., Brice, A. & Kabashi, E. Mutations in the PFN1 gene are not a common cause in patients with amyotrophic lateral sclerosis and frontotemporal lobar degeneration in France. *Neurobiol. Aging* **34**, 1709 e1–1709.e2 (2013).
50. Hadano, S. *et al.* A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nat. Genet.* **29**, 166–173 (2001).
51. Eymard-Pierre, E. *et al.* Infantile-onset ascending hereditary spastic paralysis is associated with mutations in the alsin gene. *Am. J. Hum. Genet.* **71**, 518–527 (2002).
52. Chen, Y.Z. *et al.* DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *Am. J. Hum. Genet.* **74**, 1128–1135 (2004).
53. Duquette, A. *et al.* Mutations in senataxin responsible for Quebec cluster of ataxia with neuropathy. *Ann. Neurol.* **57**, 408–414 (2005).
54. Chow, C.Y. *et al.* Deleterious variants of FIG4, a phosphoinositide phosphatase, in patients with ALS. *Am. J. Hum. Genet.* **84**, 85–88 (2009).
55. Mitchell, J. *et al.* Familial amyotrophic lateral sclerosis is associated with a mutation in D-amino acid oxidase. *Proc. Natl. Acad. Sci. USA* **107**, 7556–7561 (2010).

56. Puls, I. *et al.* Mutant dynactin in motor neuron disease. *Nat. Genet.* **33**, 455–456 (2003).
57. Farrer, M.J. *et al. DCTN1* mutations in Perry syndrome. *Nat. Genet.* **41**, 163–165 (2009).
58. Nishimura, A.L. *et al.* 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 (2004).
59. Landers, J.E. *et al.* New VAPB deletion variant and exclusion of VAPB mutations in familial ALS. *Neurology* **70**, 1179–1185 (2008).
60. Skibinski, G. *et al.* Mutations in the endosomal ESCRTIII-complex subunit CHMP2B in frontotemporal dementia. *Nat. Genet.* **37**, 806–808 (2005).
61. Parkinson, N. *et al.* ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). *Neurology* **67**, 1074–1077 (2006).
62. Cox, L.E. *et al.* Mutations in CHMP2B in lower motor neuron predominant amyotrophic lateral sclerosis (ALS). *PLoS ONE* **5**, e9872 (2010).
63. Greenway, M.J. *et al.* A novel candidate region for ALS on chromosome 14q11.2. *Neurology* **63**, 1936–1938 (2004).
64. Greenway, M.J. *et al. ANG* mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. *Nat. Genet.* **38**, 411–413 (2006).
65. Millecamps, S. *et al.* SOD1, ANG, VAPB, TARDBP, and FUS mutations in familial amyotrophic lateral sclerosis: genotype-phenotype correlations. *J. Med. Genet.* **47**, 554–560 (2010).
66. Schymick, J.C. *et al.* Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. *Lancet Neurol.* **6**, 322–328 (2007).
This was the first genome-wide association study of ALS.
67. Dunckley, T. *et al.* Whole-genome analysis of sporadic amyotrophic lateral sclerosis. *N. Engl. J. Med.* **357**, 775–788 (2007).
68. van Es, M.A. *et al.* ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. *Lancet Neurol.* **6**, 869–877 (2007).
69. Cronin, S. *et al.* A genome-wide association study of sporadic ALS in a homogenous Irish population. *Hum. Mol. Genet.* **17**, 768–774 (2008).
70. van Es, M.A. *et al.* Genetic variation in *DPP6* is associated with susceptibility to amyotrophic lateral sclerosis. *Nat. Genet.* **40**, 29–31 (2008).
71. Chiò, A. *et al.* A two-stage genome-wide association study of sporadic amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **18**, 1524–1532 (2009).
72. Landers, J.E. *et al.* Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. *Proc. Natl. Acad. Sci. USA* **106**, 9004–9009 (2009).
73. van Es, M.A. *et al.* Genome-wide association study identifies 19p13.3 (*UNC13A*) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat. Genet.* **41**, 1083–1087 (2009).
74. Traynor, B.J. *et al.* Kinesin-associated protein 3 (KIFAP3) has no effect on survival in a population-based cohort of ALS patients. *Proc. Natl. Acad. Sci. USA* **107**, 12335–12338 (2010).
75. Laaksovirta, H. *et al.* Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet Neurol.* **9**, 978–985 (2010).
This was the first genome-wide association study to identify the chromosome 9p21 association signal in the founder Finnish population.
76. Shatunov, A. *et al.* Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet Neurol.* **9**, 986–994 (2010).
77. Kwee, L.C. *et al.* A high-density genome-wide association screen of sporadic ALS in US veterans. *PLoS ONE* **7**, e32768 (2012).
78. Ahmeti, K.B. *et al.* Age of onset of amyotrophic lateral sclerosis is modulated by a locus on 1p34.1. *Neurobiol. Aging* **34**, 357 e7–357.e19 (2013).
79. Deng, M. *et al.* Genome-wide association analyses in Han Chinese identify two new susceptibility loci for amyotrophic lateral sclerosis. *Nat. Genet.* **45**, 697–700 (2013).
80. Fogh, I. *et al.* No association of DPP6 with amyotrophic lateral sclerosis in an Italian population. *Neurobiol. Aging* **32**, 966–967 (2011).
81. Daoud, H. *et al.* Analysis of the UNC13A gene as a risk factor for sporadic amyotrophic lateral sclerosis. *Arch. Neurol.* **67**, 516–517 (2010).
82. Van Hoecke, A. *et al.* EPHA4 is a disease modifier of amyotrophic lateral sclerosis in animal models and in humans. *Nat. Med.* **18**, 1418–1422 (2012).
83. Schymick, J.C., Talbot, K. & Traynor, B.J. Genetics of sporadic amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **16** (spec. no. R2): R233–R242 (2007).
84. Gwinn, K. *et al.* Amyotrophic lateral sclerosis: an emerging era of collaborative gene discovery. *PLoS ONE* **2**, e1254 (2007).
85. Hollingworth, P. *et al.* Common variants at *ABCA7*, *MS4A6A*/*MS4A4E*, *EPHA1*, *CD33* and *CD2AP* are associated with Alzheimer's disease. *Nat. Genet.* **43**, 429–435 (2011).
86. Nalls, M.A. *et al.* Imputation of sequence variants for identification of genetic risks for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet* **377**, 641–649 (2011).
87. Cronin, S., Hardiman, O. & Traynor, B.J. Ethnic variation in the incidence of ALS: a systematic review. *Neurology* **68**, 1002–1007 (2007).
88. Littler, M. & Morton, N.E. Segregation analysis of peripheral neurofibromatosis (NF1). *J. Med. Genet.* **27**, 307–310 (1990).
89. Ivanchuk, S.M., Myers, S.M., Eng, C. & Mulligan, L.M. De novo mutation of GDNF, ligand for the RET/GDNFR-alpha receptor complex, in Hirschsprung disease. *Hum. Mol. Genet.* **5**, 2023–2026 (1996).
90. Alexander, M.D. *et al.* "True" sporadic ALS associated with a novel SOD-1 mutation. *Ann. Neurol.* **52**, 680–683 (2002).
This study was the first to identify a *de novo* mutation in any ALS gene, thereby demonstrating that at least a portion of sporadic disease was due to spontaneous mutations.
91. Chiò, A. *et al.* A de novo missense mutation of the FUS gene in a "true" sporadic ALS case. *Neurobiol. Aging* **32**, 553.e23–553.e26 (2011).
92. DeJesus-Hernandez, M. *et al.* De novo truncating FUS gene mutation as a cause of sporadic amyotrophic lateral sclerosis. *Hum. Mutat.* **31**, E1377–E1389 (2010).
93. Chesi, A. *et al.* Exome sequencing to identify *de novo* mutations in sporadic ALS trios. *Nat. Neurosci.* **16**, 851–855 (2013).
This was the first study to apply exome sequencing of trios to identify *de novo* mutations as a possible cause of sporadic ALS.
94. Turner, M.R. *et al.* Controversies and priorities in amyotrophic lateral sclerosis. *Lancet Neurol.* **12**, 310–322 (2013).
95. Chiò, A. *et al.* Genetic counselling in ALS: facts, uncertainties and clinical suggestions. *J. Neurol. Neurosurg. Psychiatry* doi:10.1136/jnnp-2013-305546 (6 July 2013).
96. van Blitterswijk, M. *et al.* Evidence for an oligogenic basis of amyotrophic lateral sclerosis. *Hum. Mol. Genet.* **21**, 3776–3784 (2012).
97. Chiò, A. *et al.* ALS/FTD phenotype in two Sardinian families carrying both C9ORF72 and TARDBP mutations. *J. Neurol. Neurosurg. Psychiatry* **83**, 730–733 (2012).
98. Elden, A.C. *et al.* Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* **466**, 1069–1075 (2010).
99. Van Damme, P. *et al.* Expanded ATXN2 CAG repeat size in ALS identifies genetic overlap between ALS and SCA2. *Neurology* **76**, 2066–2072 (2011).
100. van Es, M.A. *et al.* Angiogenin variants in Parkinson disease and amyotrophic lateral sclerosis. *Ann. Neurol.* **70**, 964–973 (2011).