



King's Research Portal

DOI:

[10.1038/nrneurol.2016.182](https://doi.org/10.1038/nrneurol.2016.182)

Document Version

Peer reviewed version

[Link to publication record in King's Research Portal](#)

Citation for published version (APA):

Al-Chalabi, A., van den Berg, L. H., & Veldink, J. (2016). Gene discovery in amyotrophic lateral sclerosis: implications for clinical management. *Nature Reviews Neurology*. <https://doi.org/10.1038/nrneurol.2016.182>

Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.

Gene discovery in ALS: What's been found, what's in store, and the implications for clinical management

Abstract

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease predominantly affecting upper and lower motor neurons, leading to relentlessly progressive weakness of voluntary muscles, with death typically resulting from diaphragmatic failure within two to five years. Since the discovery of mutations in *SOD1* in 1993, which account for about 2% of ALS, there have been increasing efforts to understand the genetic component of risk in the expectation that this will reveal mechanisms causing motor neuron death, aid diagnosis and classification, and guide personalized treatments. In this Review, we outline previous and current efforts to characterize ALS genes, describe what is currently known about the genetic architecture of ALS, both in terms of the effects on family history, and the likely nature of future gene discoveries, and explore how our understanding of ALS genetics affects present and future clinical decisions. We observe that the effect of many ALS gene variants lies somewhere between mutations that greatly increase risk and common variants that have a small effect on risk, and combine this with insights from Next Generation Sequencing to explore the implications for genetic counselling.

Department of Basic and Clinical Neuroscience, Maurice Wohl Clinical Neuroscience Institute, King's College London, London SE5 9RX, UK (A. Al-Chalabi). Department of Neurology, UMC Utrecht, Netherlands (L.H. van den Berg, J.H. Veldink)

Correspondence to: A. Al-Chalabi, Maurice Wohl Clinical Neuroscience Institute, 5 Cutcombe Road, London SE5 9RX, UK

ammar.al-chalabi@kcl.ac.uk

Competing interests:

A. Al-Chalabi declares associations with the following companies: OrionPharma, Cytokinetix Inc, Mitsubishi-Tanabe Pharma, OneWorld Publications, Cold Spring Harbor Laboratory Press, serves on the Scientific Advisory Board for the ALS Association, Prize4Life and ALSGene, and serves on the Editorial Board of Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, and F1000. L. van den Berg received a grant from the Netherlands Organization for Health Research and Development (Vici scheme); serves on scientific advisory boards for Prinses Beatrix Spierfonds, Thierry Latran Foundation, Baxalta, Cytokinetix and Biogen, serves on the Editorial Board of the Journal of Neurology, Neurosurgery, and Psychiatry, Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration, and Journal of Neuromuscular Diseases. J. Veldink declares no competing interests.

Key points:

- Amyotrophic lateral sclerosis (ALS) is a syndrome resulting from many possible underlying genetic variations.
- The genetic architecture of ALS is predominantly one in which a few rare variants contribute to risk in each individual rather than a polygenic architecture in which the cumulative effect of many common variants increase risk.
- Carrying a disease mutation does not inevitably lead to ALS in every case, and many ALS genes are also implicated in other conditions, including frontotemporal dementia and cerebellar disease. The distinction between familial and sporadic ALS is not clear cut. These factors greatly complicate genetic counselling in ALS.
- The rate of gene discovery in ALS is doubling every four years.

- The data are consistent with a model in which multiple molecular steps are required to cause ALS. The causes of the steps may be genetic or environmental.

Introduction

Amyotrophic lateral sclerosis (ALS, also known as motor neuron disease) is a devastating neurodegenerative disease affecting upper and lower motor neurons and, to a variable extent, extramotor systems such as temporal circuits and behavioural and executive frontal circuits.^{1,2} An affected person becomes progressively weaker over months, until death occurs from neuromuscular respiratory failure, typically two to five years after first symptoms.

Although the peak age of onset is about 70 years, ALS can affect people of any age and is the commonest neurodegenerative disease of mid-life; the cumulative lifetime risk of ALS is about 1 in 300.³ The incidence is 1-2 per 100,000 person-years.⁴ The point prevalence however is only 5 per 100,000 persons because of the very poor prognosis.

Only a few non-genetic risk factors have been reliably confirmed for ALS, increasing age being one, although it is not clear if the risk drops off for the very elderly.⁵ Many studies also show an increased risk for males, with a 3:2 male-female ratio, but this is not true for all populations, and may depend on the age structure of the studied group, since men predominate at younger ages.⁶

Does sporadic ALS have a genetic component?

A family history of ALS or frontotemporal dementia can be obtained in a significant proportion of cases.⁷ Depending on the definition of familial ALS used, between 5% and 20% of people report a positive family history. While it may be obvious that familial ALS has a genetic component, it is not so clear whether apparently sporadic ALS has any genetic basis. Twin and other family studies have shown that the heritability of apparently sporadic ALS is about 60%, suggesting that a substantial genetic contribution is available for discovery even in those with no family history.⁸⁻¹⁰

Although making a distinction between familial and apparently sporadic (isolated) cases is useful for genetic counselling and in informing suitable research strategies for gene discovery, the boundary between the two is not clear cut. The definition of familial disease varies widely between physicians,¹¹ but even when there is a single gene variant that greatly increases the risk of ALS, the probability of obtaining a positive family history depends on the family size.¹² When the disease gene variant only contributes moderately to ALS risk or increases the risk of other conditions in addition to ALS, the probability of noting a positive family history drops further.¹³ It is therefore expected that familial ALS mutations should sometimes be found in those with apparently sporadic ALS, and this is indeed confirmed in multiple studies. The corollary is that in each person, even apparently sporadic ALS may result from a few gene variants that each confer a moderate risk, rather than the alternative scenario of the cumulative effect of multiple common gene variants each contributing a little to disease risk.

Gene discovery and ALS genetic architecture

Family based studies have been highly successful in identifying ALS genes. In the past, these used a technique in which the transmission and sharing of genetic variations within a family were used to home in on the disease gene (a method called linkage), but now it is possible to simply sequence the entire genome (whole genome sequencing), or for economy, focus on the protein coding portion of the genome where disease gene variation is likely to be found (whole exome sequencing). Even though whole genome and whole exome sequencing will miss some forms of genetic variation, the combination of these methods and linkage is an excellent method for identifying ALS genes in families.^{14,15} However, even the most frequent cause of ALS, a mutation in which a hexanucleotide repeat, GGGGCC in an intron of the *C9orf72* gene is expanded into hundreds or thousands of repeats, is not detectable by sequencing because large and repetitive sequences are missed by the technology and analysis tools currently used, emphasizing the importance of using statistical techniques, family studies and other methods in addition to whole genome sequencing.

In sporadic ALS, a widely used method for gene discovery has been to search for association of genetic variants with disease status in case-control studies. Genome-wide association studies using common variants have identified a few replicable ALS risk loci.¹⁶⁻²⁰ This method of gene discovery is based on the “common disease common variant” hypothesis under which sporadic ALS is assumed to result from the cumulative effect of multiple common genetic variants. It is therefore useful to consider whether ALS is a common disease in this context, since it will reveal the likely genetic architecture – multiple small contributions to risk or a few large contributions to risk.

Discrete traits such as schizophrenia with a lifetime risk of around 1% are sufficiently common that the common disease common variant hypothesis could apply. On the other hand, rare diseases with a genetic basis are often caused by large effect mutations in a single gene, and have a lifetime risk that is a tiny fraction of a percent. Huntington’s Disease is an example of such a condition. ALS lies somewhere between these two extremes, suggesting that it might have a genetic architecture somewhere between the two.

A review of lists of genetic risk factors for ALS (for example, <http://alsod.iop.kcl.ac.uk>)^{21,22} focussing on those in the Online Mendelian Inheritance in Man (OMIM) database²³ (Table 1), and combined with examination of genes identified from genome-wide association studies²⁴ (Table 2) allow some inferences.

ALS, like other neurodegenerative diseases, has a single, overwhelmingly large genetic signal detectable by association tests. In the case of ALS, the association is with genetic variation on chromosome 9.^{16,17} Despite the strength of this association, it was only detectable with sample sizes in the thousands, and only replicated with a further increase in sample size. The association signal corresponded to the genetic location of a signal identified in families with inherited forms of ALS.²⁵⁻³¹ We now know that both signals are

pointing to the same genetic variant, hexanucleotide repeat expansion in the *C9orf72* gene.^{32,33} The reason this expansion mutation is detectable with both case-control association studies and family based studies is because it occupies the middle ground between a rare gene variant that almost inevitably leads to disease and common gene variation that only increases risk a little (Figure 1). The hexanucleotide expansion results in a moderate increase in the risk of ALS, frontotemporal dementia or both, and is responsible for up to 10% of apparently sporadic ALS in some populations, and about 30% of familial ALS.

A few common gene variants are replicably associated with ALS, increasing risk by a small amount only. One is in the *UNC13A* gene.^{34,35} Expansion of the CAG trinucleotide repeat in the *ATXN2* gene is known to cause spinocerebellar ataxia, but if the repeats are of intermediate size, smaller than the range associated with spinocerebellar ataxia, they are a well replicated risk factor for ALS.³⁶ Common variants in the *MOBP* gene and the *SCFD1* gene, also show replicated association with ALS (Table 2). Variants in *ELP3* have been associated with ALS, and although, because of the nature of the variation, no replication study has been done, there is functional work to support the findings.³⁷ Similarly, *SMN1* copy number variation,³⁸ and indel mutation in the *NEFH* gene are considered risk factors (Tables 1 and 2).³⁹

Single nucleotide variations account for only 8.5% of the heritability of ALS, which is much smaller than the equivalent for schizophrenia for example at 21%, suggesting that rare variation, structural variants such as large deletions or inversions, and repeat sequences that are not captured by current high throughput techniques must account for much of the heritability.^{19,40} In support of this view, a large genome-wide association study of ALS has shown a larger than expected contribution from low-frequency variants in genetic susceptibility to ALS (See Box 1).

Sequencing as a gene finding strategy

The disproportionately large role for low-frequency variation in the genetic architecture of ALS, including apparently sporadic ALS, is a strong argument to look for such variants. When there is no family history, a case-control design is the simplest approach, but requires the ability to sequence the whole genome to assay the rare variation. Although whole genome sequencing is now feasible, the major problem is that it is not straightforward to interpret findings. One might expect that a protein changing mutation found in patients but not controls must indicate pathogenicity, but this is not the case, since many protein changing mutations occur in the general population, as can be seen from 60,706 sequenced control exomes in the ExAC database, <http://exac.broadinstitute.org>.⁴¹ On average, each person has ten major protein truncating or extending mutations without obvious disease effects. Furthermore, in a late onset disease one would expect as yet unaffected controls to carry disease-causing mutations, and when changes are rare, the numbers available for statistical testing are small. Since rare variation is also likely to be specific to certain populations, replication studies are difficult and this leads to uncertainty around the statistical declaration of a relationship between a variant and disease. One potential test is post mortem identification of protein encoded by a putative ALS gene in pathological inclusions or aggregates, but this is only useful if such material is available and if the protein is found.⁴²

These problems are illustrated with *SOD1* mutation in ALS. The initial linkage based studies showing a disease associated locus on chromosome 21 led to the identification of the p.A5V mutation of *SOD1* as a causative variant.^{43,44} Subsequently, other *SOD1* mutations were identified, with either linkage or functional evidence or both. However, the discovery of *SOD1* as a pathogenic ALS gene led to widespread sequencing of the gene in families and those with apparently sporadic ALS, leading ultimately to more than 130 mutations being identified in this 153 amino acid gene (<http://alsod.iop.kcl.ac.uk>). Identification of a protein changing mutation in a known ALS gene in a patient with ALS naturally leads to the assumption that the mutation is causative, but functional evidence or evidence for familial segregation is often lacking or limited.^{13,45} A similar problem now arises in the large scale sequencing of

genomes of people with ALS. Identification of a rare variant in someone with ALS cannot on its own be regarded as evidence of pathogenicity, even if it occurs in a known ALS gene. Because such variants are by definition rare, it is not possible to provide even statistical support through an association with ALS. Using bioinformatics methods to predict if a variant is detrimental is not effective either, as many of the protein coding variants in the Exome Aggregation Consortium (ExAC) database are predicted to have detrimental effects, even in *SOD1*, despite being found in a general population sample.

The problem of interpretation of findings is even more difficult for the 99% of the genome that is non-protein coding. Although it is a popular belief that we know the genetic code, this is only true for the protein coding portion of the genome. For the non-coding component we have made some progress⁴⁶ but we do not completely understand how the nuclear and cellular machinery interprets the sequence. It is therefore very difficult to understand whether a variant is pathogenic when found outside the coding regions. There are also further confounders that we are only now beginning to understand, such as the 3-dimensional conformation of the genome,⁴⁷ the interaction of the genome structure with transcription⁴⁸ and somatic mosaicism in which different cells within a person have their own mutations collected through mitosis during development.⁴⁹

One response to this problem is collaboration on a truly global scale. This strategy allows sufficient numbers to be assayed that the less rare variants may be seen in more than one individual, allowing statistical support, or may be seen to cluster in a particular domain allowing functional studies to show pathogenicity of lesions in that region of a protein. If those interpretations are not possible, burden testing can be used in which a simple count of rare variants can be taken per gene in cases and controls, with statistical excess in one group used to support an argument for a pathological role of the relevant gene. One such initiative is Project MinE (<http://www.projectmine.com>), an international whole genome sequencing consortium using crowdsourcing for funding, which is on target to achieve a goal

of 15,000 ALS whole genomes and 7,500 control sequences. Even this project will need to work with data from other populations to increase the number of controls available for example.

Implications of genetic findings for the development of new treatments

Recent efforts using family based whole exome and whole genome sequencing, and large-scale genome-wide case-control association studies in ALS have already uncovered a number of novel ALS risk variants in various genes and illuminated potential molecular pathways (Table 1). These include an increased burden of protein-changing and loss of function mutations in the genes *TBK1*, *NEK1*, and a gene coding for a mitochondrial protein, *C21orf2*.¹⁹ The *NEK1* protein interacts with *C21orf2* as well as other proteins involved in motor neuron degeneration, including *ALS2* and *VAPB*. *TBK1* phosphorylates a protein implicated in ALS, *OPTN*.⁵⁰ *TBK1* mutations also appear to segregate with disease in pedigrees with ALS and FTD.¹⁵ *TBK1* is known to be involved in autophagy, especially autophagosome maturation as well as the clearance of pathological aggregates through the proteasome. Other proteins involved in the proteasomal pathway include *UNC13A*, *ATXN2*, *UBQLN2*, *SQSTM1*, and *SARM1*. Through the NF-kappaB pathway, *TBK1* also has a role in innate immunity signaling, which is related to neuroinflammation and may have a role in risk and rate of disease progression in ALS.⁵¹

C21orf2 is a poorly characterized protein, but through its interaction with *NEK1* it has been shown to be crucial for proper DNA repair.⁵² Both *NEK1* and *C21orf2* are part of the “ciliome” and are required for the formation and maintenance of primary cilia.⁵³ Defects in primary cilia are associated with various neurological disorders and cilia numbers are decreased in G93A *SOD1* transgenic mice.⁵⁴ Microtubule organization and kinesin/dynein intra-flagellar transport are essential to maintain cilia structure and function, and it is known that disruption of the microtubule cytoskeleton is associated with the development of ALS,⁵⁵ and mutations of the dynein subunit dynactin (*DCTN1*) are a rare cause of familial ALS.⁵⁶ Other cytoskeletal

proteins are also implicated in ALS through genetics. These include VAPB, VCP, SCFD1, OPTN, PFN1, NEFH, and TUBA4A.

Previously identified ALS genes with varying levels of support, code for proteins involved in RNA processing (TDP43, FUS, SETX, ELP3, ANG, TAF15 and others). RNA processing is an ubiquitous process, and it is not yet clear why such defects might result in specific injury to motor neurons.

Thus, molecular pathways relevant to ALS are emerging, and include DNA repair, RNA processing, autophagy, inflammation, protein degradation, mitochondrial dysfunction and cytoskeletal organisation (Table 1). These are all logical therapeutic targets, but the implication of new findings for individual patients may well be difficult to interpret. Progress in the discovery of novel ALS related genes is greatly accelerating (Figure 2). Biological insights will grow concomitantly and therefore fuel novel therapeutic developments. Also, these discoveries are paving the way for precision medicine in ALS through the precise knock-down or even gene-editing of specific ALS associated mutations (Box 2).

Implications of genetic findings for counselling in ALS

Compounding the difficulty interpreting newly identified mutations are three genetic effects, oligogenic inheritance, pleiotropy and reduced penetrance. Oligogenic inheritance is when a single mutation is not sufficient to cause disease despite significantly increasing risk. Other factors are required such as other gene variants, to cause ALS. This was first described in a French family in which affected individuals had two different mutations, one in the maternal and the other in the paternal copy of the *SOD1* gene,⁵⁷ but more recently has been seen in affected individuals carrying combinations of risk variants in *FUS*, *TARDBP*, *C9orf72*, *SOD1*, *VAPB*, *OPTN1*, and *ANG*.⁵⁸⁻⁶⁰ *C9orf72* hexanucleotide repeat expansion confers moderate risk, not as large as for typical familial disease genes, but far greater than the modest odds ratios seen for common variants associated with ALS. If oligogenic inheritance is a frequent

theme in ALS, all the relevant gene variants will need to be identified and tracked through a family to reveal risk and will greatly complicate the interpretation of a positive gene test for genetic counsellors. An important and recent finding that further exemplifies how oligogenic inheritance might affect genetic counselling in the future is the finding that knockdown of the *SUPT4H1* gene greatly reduces expression of the *C9orf72* hexanucleotide repeat expansion.⁶¹ Deletions or loss of function mutations in *SUPT4H1*, therefore, might be a natural modifier of *C9orf72* mediated toxicity and understanding the genetic variants an individual carries in *SUPT4H1* would then be essential in interpreting the effect of being a *C9orf72* hexanucleotide expansion mutation carrier.

Pleiotropy is the observation that a particular gene mutation may result in different diseases, either simultaneously or in different individuals. For example, expansion mutation of *C9orf72* can result in ALS, frontotemporal dementia or both (Figure 3).⁶² For *ATXN2* the situation is complicated further because the exact variation influences the disease: those with up to 28 CAG trinucleotide repeats are normal, those with 29 to 32 repeats at risk of ALS, and those with 33 or more repeats at risk of spinocerebellar ataxia, with little overlap at the boundaries.³⁶ The repeat sizes are not stable between generations. This phenomenon of pleiotropy in ALS is increasingly recognised, and extends most frequently to frontotemporal dementia, but also to ataxia, parkinsonism, mitochondrial myopathies, Paget's disease, Alzheimer-type dementia, psychiatric disorders such as schizophrenia, and possibly multiple sclerosis, associated with variants in *C9orf72*, *ATXN2*, *TBK1*, *FUS*, *C21orf2*, *NEK1*, *MATR3*, *CHCHD10*, *VCP*, *hnRNPA1*, *hnRNPA2B1* and others (Table 1).^{7,13,19,20} The implications for genetic counselling are that the family history may be incomplete, since the different diseases are not correctly recognised as a positive family history, and the interpretation of a positive gene test for other family members is no longer limited to the risk of developing a single condition.

A related genetic phenomenon is age dependent penetrance. In this context, penetrance is the probability of developing a disease if a mutation carrier. All ALS genes and many genes for related conditions show age dependent penetrance, with the risk of manifesting disease increasing with age. This means that development of a disease is not inevitable just because someone carries a risk variant, since the age at which disease manifests may be older than the lifespan of the person. From a clinical perspective, this leads to the disease skipping generations and therefore impacts the likelihood of a positive family history, and also means that the reduced risk of developing a disease needs to be explained to gene carriers or those at risk, even though the exact profile of risk reduction is complex or unknown. There are also ethical difficulties for prenatal screening, for example, termination of a pregnancy for a fetus carrying *C9orf72* expansion. These are very complex matters and counselling should be provided by trained clinical geneticists.⁶³

This complexity is perhaps best illustrated by further considering the hexanucleotide repeat expansion mutation of *C9orf72*, which is carried by up to 10% of all people with ALS in some populations, regardless of family history.⁶⁴ A significant problem therefore, is whether everyone, even those with apparently sporadic ALS, should be tested. The lack of a family history of ALS is not strong evidence against this single gene cause. On the other hand, the correct interpretation of a positive result is not clear, since the mutation may not result in ALS in the offspring or relatives, and may not cause disease at all. Furthermore, at present, no treatment is possible, although that may change as genetic therapies become available. The correct approach is still under debate and there are differences between countries with some screening all patients and others only if there is a family history of ALS in a first degree relative.

Multistep model

These three genetic phenomena can be explained through a multistep model of ALS pathogenesis, which has recently been shown to fit the incidence profile of the disease.⁶⁵

The evidence fits a model showing that on average, ALS results from six pathological steps which may themselves result from one or more genetic or environmental risk factors. Also, the disproportionately large role for low-frequency high risk variation in the genetic architecture of ALS as opposed to the concerted action of thousands of low risk variants, is consistent with this model. The multistep model also explains gene x environment interaction, since some of the steps would be triggered by genes and others by environmental factors, but because these occur within a specific pathway, the environmental trigger is only relevant within the context of the genetic trigger.

Implications of genetics for diagnosis and prognosis

A desirable scenario is that knowledge of the gene profile of an affected individual provides a sensitive and specific diagnostic test for ALS. Indeed, it has been proposed that the El Escorial criteria should have provision for diagnosis of ALS based on identification of mutation in a familial ALS gene.⁶⁶ However, because of the difficulties in interpreting the meaning of a mutation when there is no family history, there are challenges in using such an approach for most people with ALS. A gene profile might still be useful in allowing classification into a subtype suitable for targeting with a specific treatment strategy, either through gene therapy, or because a specific pathway can be targeted.

To aid with interpretation and genetic counselling, we have provided information in Table 1. The genes where there is strong evidence for “genic constraint” (<http://exac.broadinstitute.org>), i.e. with “good” to “fairly good” Interpretability of genetic findings for counselling, are most suitable to routinely test if the phenotype allows for it. If genetic results are returned, even previously unknown mutations that lead to a truncated or absent protein can be regarded as being pathogenic. The “low” category is typically the category of genes that can have amino acid changing mutations which are probably pathogenic but not necessarily always, given the frequency of observed mutations in the general population such as those in the ExAC database, and the lack of evidence for

segregation of genetic variation with disease within a family. One should be aware of this when routinely testing these genes in the clinic. The Variants of Uncertain Significance (VUS) category genes should really not be routinely tested in clinic since they encode proteins that are apparently highly tolerant of amino acid changing mutations.

Risk profiling in healthy individuals using methods to evaluate the total effect of genetic risk is possible.⁶⁷ Using individual gene profiles to predict the development of ALS by screening the population is not practical however, and is unlikely ever to be so. The major problem is the risk of false positive tests in a rare condition that cannot be prevented or avoided. Even in ideal circumstances, it is likely that screening would do more harm than good.

There is a role for genetics in predicting prognosis. At present this is restricted to very basic observations such as a better or worse outlook being likely in those with certain mutations of *SOD1* or an increased risk of frontotemporal dementia in those with *C9orf72* expansion mutation. However, association studies examining survival offer the opportunity to generate a prognostic genetic score that could be used to stratify in clinical trials, or, in combination with clinical features, be used to inform clinical care. Already, there is replicated association of variants in the *UNC13A* gene and initial association of variants in the *CAMTA1* gene with worse prognosis. Such findings are potentially important therapeutically, since the mechanism of ALS causation and the mechanism of disease progression may well be different.

Conclusions

Genetic studies of ALS are at an exciting and crucial phase in which advances in technology and unprecedented large scale international collaboration are combining to rapidly increase our understanding of the causes of this disease.

Acknowledgements:

This is an EU Joint Programme - Neurodegenerative Disease Research (JPND) project (STRENGTH, SOPHIA, ALS-Care). The project is supported through the following funding organisations under the aegis of JPND - www.jpnd.eu (United Kingdom, Medical Research Council and Economic and Social Research Council; Netherlands, ZonMW). AAC receives salary support from the National Institute for Health Research (NIHR) Dementia Biomedical Research Unit at South London and Maudsley NHS Foundation Trust and King's College London. The work leading up to this publication was funded by the European Community's Health Seventh Framework Programme (FP7/2007–2013; grant agreement number 259867) and Horizon 2020 Programme (H2020-PHC-2014-two-stage; grant agreement number 633413). This study was supported by ZonMW under the frame of E-Rare-2, the ERA Net for Research on Rare Diseases (PYRAMID).

BOX1

Types of genetic studies

Genetic studies come in many forms and flavours depending on the genetic architecture of the disease or trait of interest. Below is an overview of different study designs.

Genetic linkage studies

A family study design based on the phenomenon whereby alleles at different loci are transmitted together from parents to offspring more often than expected by chance in relation to a disease. Linkage studies are typically good at identifying disease variants that are in themselves sufficient to cause disease (for example in Huntington's chorea), but not so good at finding variants that only increase risk a little.

Candidate gene studies

The selection of one or more candidate genes based on a biologically plausible hypothesis, in order to compare genetic variation between cases and controls. Most past candidate gene studies have not replicated robustly.

Genome-wide association studies (GWAS).

GWAS either have a case-control design with disease status or a quantitative trait as outcome (e.g. blood pressure). Typically hundreds of thousands of common variants are genotyped simultaneously using a DNA microarray, and frequencies of different variants are compared between cases and controls or correlated with the trait of interest. GWAS are well suited to find common variants that increase risk for a disease but are not necessarily causative.

Genome-wide next generation sequencing studies

Whole genome sequencing based studies typically include both common and rare genetic variation. Because rare variants can so infrequent that statistical tests are not reliable, special techniques are needed. One method takes the aggregate of rare variation in a specific gene in cases and compares it with the aggregate in controls. Gene sizes vary, and large genes have more chance to accumulate rare variation, so some tests are weighted to account for these biases. Another method tests each variant independently, and then combines the results across multiple DNA sequences to test for association. This allows for the fact that some variants might be protective and otherwise cancel out risk variants. . Examples of such tests are SKAT, C-alpha, and EREC.

END BOX

BOX2

From risk genes to precision medicine

The recent notion that the bulk of the genetic risk factors that remain to be identified in ALS are likely to be rare variants with intermediate to large effects on risk has important consequences for future drug development. There is a clear need to accelerate the development of new ALS drugs since many clinical trials have been negative in the past. The hope is that the discovery of ALS risk genes will help this process.

In general, three strategies are used to identify novel ALS risk genes allowing us to arrive at a precision medicine state.

1. The identification of ALS risk genes that act in specific molecular pathways may allow for a stratified treatment approach using compounds that target the pathway. ALS genes do not appear to all act through the same mechanism. For example, *FUS* and *TARDBP* appear to be mainly active in RNA metabolism, *TUBA4A* and *PFN1* in axonal and cytoskeletal biology, *VCP*, *OPTN* and *TBK1* in autophagy, and *UBQLN2* and others in protein stability, conformation and degradation. Many other pathways will follow, and it is highly plausible that different compounds will be needed dependent on the class of ALS genes that is involved in a subgroup of patients.
2. The identification of specific mutations that act through a toxic-gain of function in ALS may offer an opportunity for more specific precision medicine. Most notable examples are *SOD1* and *C9orf72*. The first steps towards this approach in ALS have already been taken, as a successful phase 1 study with *SOD1* antisense therapy has already been performed and a phase 2/3 trial is under way.⁶⁸ Also, many research groups are working on the development of gene-targeted therapies through antisense oligonucleotides, and viral delivery of si-RNA, in particular for *C9orf72* mutation. One special caveat here is that overall knock-down of ALS gene

expression, affecting both the wild-type and variant allele, might also have detrimental effects. This, for example appears to be the case in C9orf72.^{69,70}

3. The ultimate form of precision medicine, but also the most distant one, is that of genome editing. Recent exciting breakthroughs in molecular biology have made it possible to induce mutations or repair mutations through a biological machinery originally discovered in bacteria called CRISPR/cas9.^{71,72} This form of precision medicine is in fact “pinpoint-medicine”, meaning that within one ALS gene, every damaging mutation involved would need its own specific treatment. Current elegant examples are the elaborate efforts in Duchenne muscular dystrophy where specific antisense therapies are being tested to induce exon skipping to improve dystrophin levels in muscle and thereby improved functional outcome in patients. Since dystrophin is a very large gene with 79 exons and many mutations have been described, many specific antisense oligomers will be needed that each target the disease in a very small subset of patients.

For some ALS mutations it is very clear that they are directly disease causing and therefore amenable to targeting with a precision medicine approach. Nevertheless, while mutations or genes can have robust statistical association with ALS, for many mutations direct pathogenicity has not been demonstrated. The hope is that high-throughput screening in neuronal cell models, for example, based on patient-derived induced pluripotent stem cells, will make this process easier. Importantly, these functional assays should be done only when there is sound genetic evidence to begin with.⁷³

END BOX

Definitions

Locus

A chromosomal region, often defined by a property such as coding for protein or RNA.

Allele

A genetic variant

Recombination

Two genetic variants can be inherited from one parent but originate from different grandparents. If such variants are on the same chromosome, this means that sections of chromosomes have swapped during meiosis, a process called recombination.

Linkage disequilibrium

A measure of whether gene variants are associated with each other. Variants that are in linkage disequilibrium are found together on the same haplotype more often than expected by chance.

Haplotypes

Combinations of genetic variants that are inherited together.

Penetrance

The conditional probability of a phenotype (for example ALS) given a genotype.

Genetic pleiotropy

The situation where genetic variants can lead to more than one disease or trait. The diseases may appear unrelated from a clinical viewpoint. Decades ago, no one would have grouped ALS with FTD, but although there was increasing evidence for a clinical overlap over the last twenty years, the discovery of the *C9orf72* repeat expansion in ALS-FTD, has

dramatically confirmed an aetiological and pathological overlap. The *C9orf72* genetic variant is also associated with Parkinsonism, Huntington's chorea, Alzheimer's disease, psychosis and bipolar disorder.

Heritability

The proportion of phenotypic variation in a population that is attributable to genetic variation among individuals.

Genetic architecture

The number of risk variants underlying disease, their relative frequencies, the size of their effects on risk and their mode of interaction.

Next-generation sequencing (NGS)

Highly parallel DNA-sequencing technologies that produce many hundreds of thousands or millions of short reads of DNA (25–500 bp) for a low cost and in a short time. The reads need to be assembled into a full genome by supercomputer.

Mosaic mutations

Mutations that are present in only a proportion of cells in the body.

Structural variation

Occurs in DNA regions generally greater than 1 kilobase in size, and includes genomic imbalances (namely, insertions and deletions, also known as copy number variants), inversions and translocations.

De novo mutations

Non-inherited novel mutations in an individual that result from a germline mutation.

END Definitions

Figures

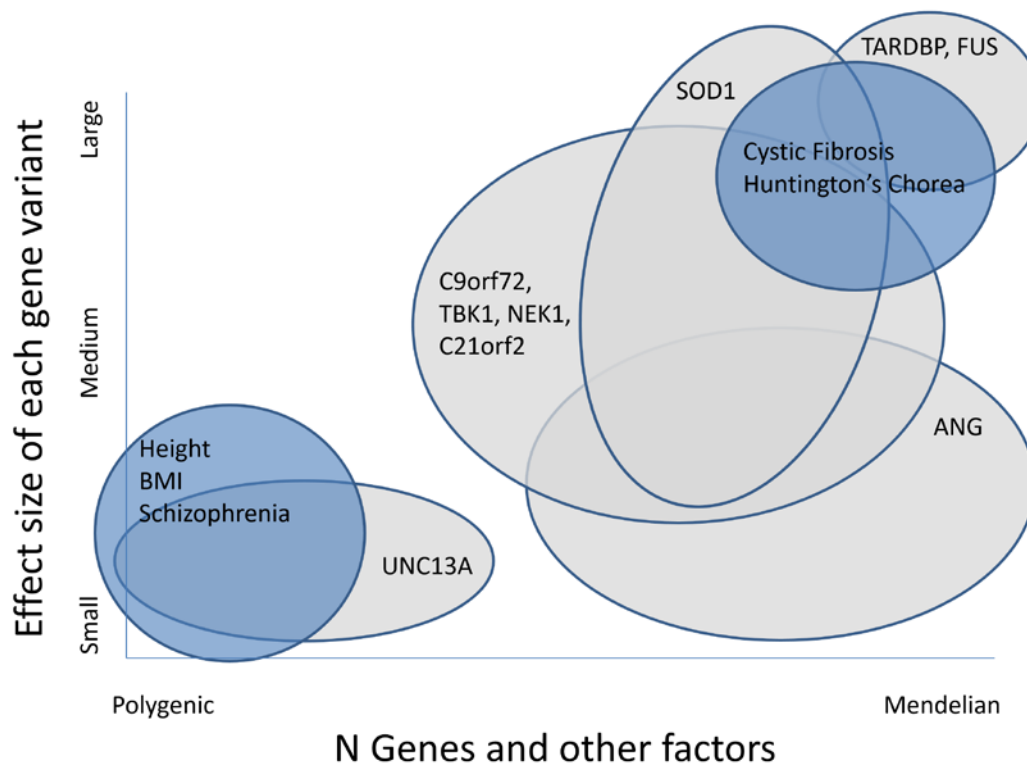


Figure 1. The relationship between allele frequency and effect size for ALS genes

Traits such as height, body mass index (BMI) and schizophrenia are influenced by the cumulative effect of tens or hundreds of gene variants, each only contributing a little.

Because of the small effect of each variant, there is only weak removal from the population

by natural selection, and they can reach high frequency, becoming common. Diseases such as cystic fibrosis or Huntington's chorea result from single gene mutations of very large effect, greatly increasing the risk of disease. Because of the large effect, such variants tend to be removed by natural selection and remain rare in the population, unless (as is the case for cystic fibrosis) they confer some selective advantage in certain environments. ALS has examples of single, large effect genes and small effect genes, but the majority of variants have an effect size somewhere in between.

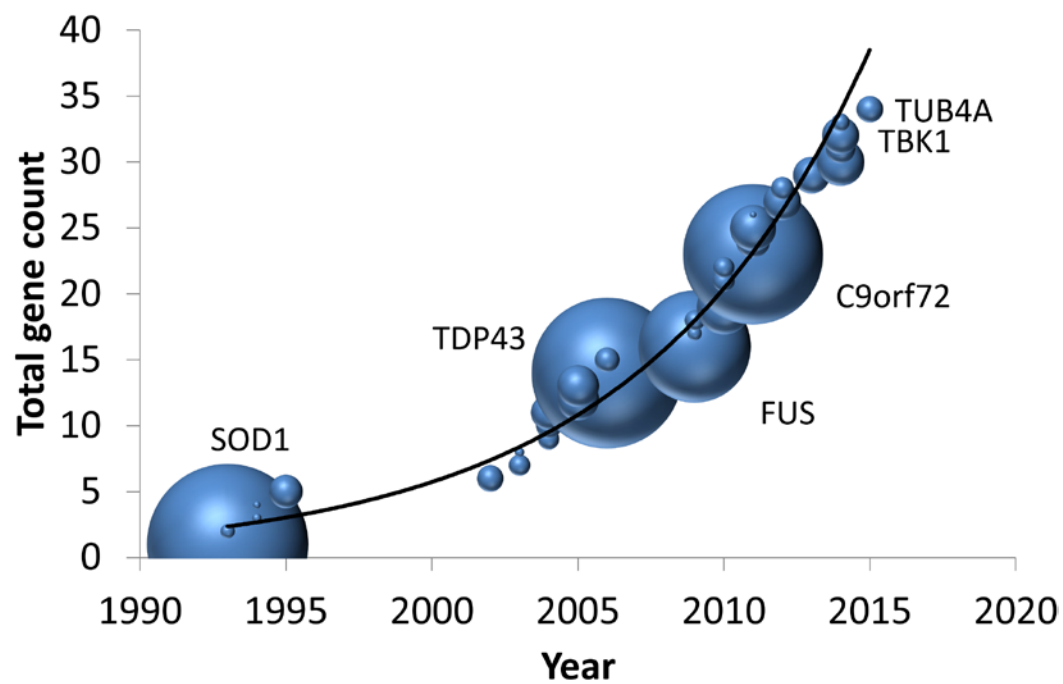


Figure 2. Gene count by year of publication

The size of each sphere is in direct proportion to the number of ALS related publications for that gene. Gene count is doubling every four years. TDP43 is coded by the TARDBP gene.

Data kindly provided by Dr William Sproviero.

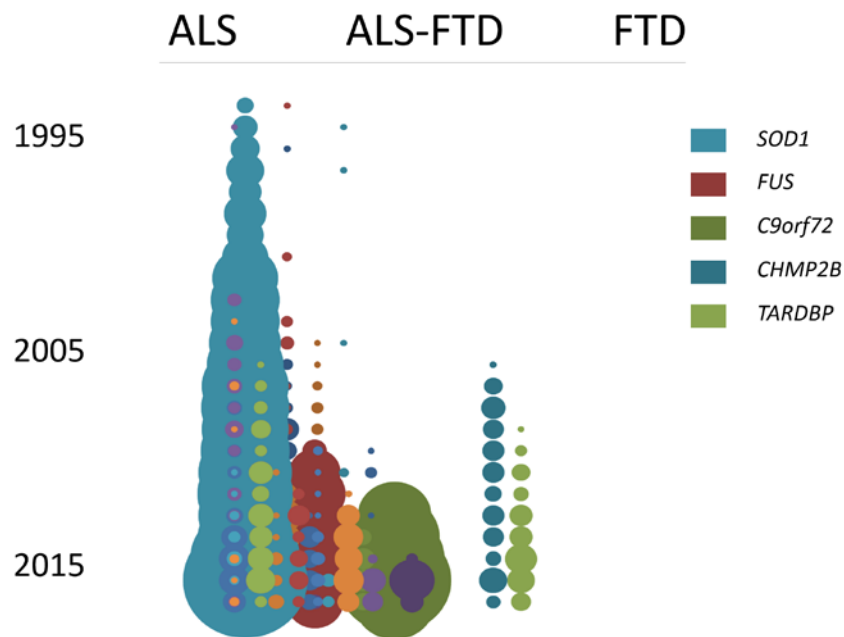


Figure 3. A stalagmite plot showing genetic pleiotropy in ALS

The size of each plotted point is in direct proportion to the number of ALS and frontotemporal dementia (FTD) publications referencing that gene, with year on the Y axis. Each gene corresponds to a different colour stalagmite (key for the most important from left to right shown). The position on the X-axis corresponds to the proportion of publications in ALS vs FTD.

Table 1. ALS genes as listed in OMIM, supplemented with recent discoveries

Locus	Name	Gene/Locus	Phenotype	Inheritance	Penetrance	Other clinical features	Ease of interpretation for counselling	Initial genetic evidence	Freq in ALS*	Biological processes	Evidence for genic constraint	Refs
1p36.22	ALS 10, with or without FTD	TARDBP (encoding TDP-43)	ALS; ALS-FTD	Dominant (recessive rare)	Can be incomplete, mostly complete	Supranuclear palsy, chorea, FTD	Good	Candidate gene and linkage	1%	DNA/RNA metabolism	High pLI	74-76
2p13.1	{ALS, susceptibility to}	DCTN1, HMN7B	ALS	Risk gene	Incomplete		Low	Candidate gene in case control study	UNK	Vesicle trafficking	Low pLI but low pNull	77
2q33.1	ALS 2, juvenile	ALS2, ALSJ, PLSJ, IAHSP	UMN predominant; juvenile	Recessive	Complete		Fairly good but note atypical phenotype	Linkage	<1%	Endosomal dynamics	Low pLI but low pNull	78,79
2q34	ALS 19	ERBB4, HER4, ALS19	ALS	Dominant	Can be incomplete, mostly complete		Good	Linkage	UNK	Neuronal development; synaptic plasticity	High pLI	80
2q35	ALS 22 with or without FTD	TUBA4A, TUBA1, ALS22	ALS; ALS-FTD	Dominant	UNK		Fairly good	Whole exome burden	UNK	Cytoskeleton architecture and dynamics	High pRec	55
3p11.2	ALS 17	CHMP2B, DMT1, VPS2B, ALS17	ALS	Risk gene	UNK	FTD	Fairly good	Candidate gene in case control study	UNK	Autophagy; lysosomal pathway	High pRec	81
4q33	To be assigned	NEK1	ALS	Risk gene	Incomplete	Short rib-polydactyly syndrome; renal pathology	Low	Homozygosity mapping followed by candidate gene sequencing; whole exome burden	3%	DNA repair; cytoskeleton architecture and dynamics	Low pLI but low pNull	20,50,82

5q31.2	ALS 21	MATR3, MPD2, ALS21	Brisk reflexes and split hand phenomenon LMN predominant disease	Dominant	UNK	(Distal) myopathy, vocal cord and pharyngeal weakness, FTD	Good but note atypical phenotype	Whole exome filtering	UNK	DNA/RNA metabolism	High pLI	83,84
5q35.3	FTD and/or ALS 3	SQSTM1, P62, PDB3, FTDALS3	ALS; ALS-FTD	Dominant	UNK	Paget's disease, FTD	Fairly good	Candidate gene in case control and pedigrees with evidence for segregation	UNK	Autophagy	High pRec	85-88
6q21	ALS 11	FIG4, KIAA0274, SAC3, ALS11, YVS, BTOP	ALS; PLS	Risk gene	UNK	CMT4J	VUS	Candidate gene in case control study	UNK	Unknown	No	89
9p21.2	FTD and/or ALS 1	C9orf72, FTDALS1, FTDALS, ALSFTD	ALS; ALS-FTD	Dominant	Can be incomplete	Parkinsonism, Huntington phenocopies, Alzheimer's disease, schizophrenia, psychosis and bipolar disorder	Fairly good	Linkage and chromosome 9 sequencing	10%	Toxic RNA species, loss of protein or toxic repeat dipeptides aggregation	NA	32,33
9p13.3	ALS 16, juvenile	SIGMAR1, SRBP, ALS16, DSMA2	Juvenile ALS	Recessive	UNK		VUS	Homozygosity mapping followed by candidate gene sequencing	UNK	Endoplasmic reticulum chaperone	No	90
9p13.3	ALS 14, with or without FTD	VCP, IBMPFD1, ALS14, CMT2Y	ALS; ALS-FTD. LMN predominant disease	Dominant	UNK	Inclusion body myopathy with early-onset Paget disease and frontotemporal dementia (IBMPFD)	Good but note phenotype	Whole exome filtering	UNK	Autophagy	High pLI	91

9q34.11	To be assigned	GLE1	ALS	Risk gene; not yet clear	Incomplete		Fairly good	Candidate gene in case control study + evidence for segregation in one pedigree	UNK	RNA export mediator	High pRec	92
9q34.13	ALS 4, juvenile	SETX, SCAR1, AOA2, ALS4	Juvenile ALS	Dominant	UNK	AOA2, cerebellar ataxia, distal motor neuropathy (with brisk reflexes)	Low	Linkage	UNK	DNA/RNA processing	Low pLI but low pNull	92-96
10p13	ALS 12	OPTN, GLC1E, FIP2, HYPL, NRP, ALS12	ALS	Recessive and dominant	UNK	Primary open angle glaucoma, FTD	Fairly good	Homozygosity mapping followed by candidate gene sequencing	UNK	Autophagy	High pRec	97,98
12q13.12	{ALS, susceptibility to}	PRPH	ALS	Risk gene	UNK		VUS	Candidate gene in case control study	UNK	Axonal regrowth	No	99
12q13.13	ALS 20	HNRNPA1, IBMPFD3, ALS20	IBMPFD/ALS	Dominant	Can be incomplete, mostly complete	IBMPFD	Good but note atypical phenotype	Linkage and exome sequencing	UNK	RNA metabolism	High pLI	100
12q14.2	FTD and/or ALS 4	TBK1, NAK, FTDALS4	ALS; ALS-FTD	Risk gene; Dominant	Incomplete, can be complete	FTD	Good	Linkage and exome sequencing and whole exome burden testing	1%	Autophagy; neuroinflammation	High pLI	15,50,98
12q24.12	{ALS, susceptibility to, 13}	ATXN2, ATX2, SCA2, ASL13	ALS	Dominant (recessive rare)	Incomplete	Longer repeat sizes: spinocerebellar ataxia, parkinsonism	Fairly good	Candidate gene in case control study and pedigree with	1-2%	RNA metabolism	NA	36,101,102

								evidence for segregation				
14q11.2	ALS 9	ANG, RNASE5, ALS9	ALS, more bulbar	Risk gene	Incomplete	ANG mutations described together with FUS and TARDBP mutations	VUS	Candidate gene in case control study	<1%	Blood vessel formation; anti-immunity	No	58,103-105
15q21.1	ALS 5, juvenile	SPG11 (SPATACSI N)	Juvenile ALS	Recessive	Complete	HSP-11, CMT2X	Fairly good	Linkage	UNK	DNA repair	High pRec	106,107
16p11.2	ALS 6, with or without FTD	FUS, TLS, ALS6, ETM4	ALS; ALS-FTD. May be young onset, aggressive ALS, especially mutation p.R521C. LMN predominant disease	Dominant (recessive rare), de novo	Can be incomplete, mostly complete	Hereditary essential tremor-4, FTD.	Good	Linkage with candidate gene approach	1%	DNA/RNA metabolism	High pLI	108-110
16p13.3	To be assigned	CCNF16p13.3-p12.3	ALS; PLS	Dominant	Complete		Low	Linkage and exome sequencing	UNK	Autophagy	Low pLI but low pNull	14
17p13.2	ALS 18	PFN1, ALS18	ALS	Dominant	Can be incomplete, mostly complete		Low	Whole exome filtering	UNK	Cytoskeleton architecture and dynamics	Low pLI but low pNull	111
18q21	ALS 3	ALS3	ALS	Dominant	UNK			Linkage	UNK	Unknown	NA	112
20p13	ALS 7	ALS7	ALS	Dominant	UNK			Linkage	UNK	Unknown	NA	113
20q13.32	ALS 8	VAPB, VAPC, ALS8	Slowly progressive ALS; LMN predominant disease	Dominant	Complete (one family)	Essential tremor	Low	Linkage	<1%	Vesicle trafficking	Low pLI but low pNull	114

21q22.3	To be assigned	C21orf2	ALS	Risk gene	Incomplete		VUS	GWAS with custom reference panel imputation	2%	DNA repair; cytoskeleton architecture and dynamics	No	19
21q22.11	ALS 1	SOD1, ALS1	ALS; LMN predominant disease	Dominant (recessive rare), de novo	Can be incomplete	Cerebellar ataxia and autonomic dysfunction (rare), FTD (rare)	Fairly good	Linkage	1-2%	Autophagy; toxic aggregation	Low pLI but low pNull	43,115,116
22q11.23	FTD and/or ALS 2	CHCHD10, FTDALS2, SMAJ, IMMD	ALS with myopathy, ataxia and FTD. LMN predominant disease	Dominant	Can be incomplete, mostly complete	Parkinsonism	VUS	Whole exome filtering	UNK**	Mitochondrial function	No	117
22q12.2	{ALS, susceptibility to}	NEFH, CMT2CC	ALS	Risk gene	Incomplete		Low	Candidate gene in case control study	UNK	Axonal transport; cytoskeleton architecture and dynamics	High pRec	39,118-121
Xp11.21	ALS 15, with or without FTD	UBQLN2, PLIC2, CHAP1, ALS15	ALS; ALS-FTD. UMN predominant disease	X-linked dominant	Incomplete	Can be juvenile	Low	Linkage	UNK	Autophagy	Low pLI but low pNull	122,123

Table 1 Legend:

The listed genes are biased towards European populations since there is limited genetic evidence for other populations.

Interpretability for genetic counselling is based on genic constraint scores,⁴¹ phenotype and level of genetic evidence for involvement in ALS. Good means that interpretation is straightforward in most cases. Fairly good means that in many cases it may be possible to determine if the variant found is relevant to ALS. Low means that interpretation may be difficult because the encoded protein is tolerant to a degree of loss of function, and there may also be reduced

penetrance, or there may be limited evidence for involvement in ALS. VUS indicates that variants of uncertain significance will be frequent because the protein encoded by the gene is tolerant to loss of function. In the VUS category, one will find many missense or even loss-of-function mutations that can be found in the general population. One should be cautious, therefore, to do routine testing on VUS category genes in ALS patients.

Genic constraint interpretation:

No: (No constraint). These are genes that are very tolerant to missense or loss of function mutations (resulting in truncated or absent protein) in the general population.

high pLI: These are genes that are highly intolerant to any loss of function mutations, either heterozygous or homozygous. They are near Mendelian genes, in which mutations have very high or full penetrance.

high pRec: These genes are mostly intolerant of homozygous loss of function mutations, but generally tolerant of heterozygous mutations.

low pLI but low pNull: These genes are intolerant to some heterozygous and some homozygous mutations, but tolerant to others

UMN: Upper motor neuron; LMN: Lower motor neuron; UNK: Unknown; Freq: Frequency; Refs: References

* Assuming a rate of familial ALS of 10% in all ALS and based on European populations

** Many variants have been published but any damaging effect of many of those in ALS and or FTD is still unclear

Table 2. Replicated genome-wide association findings in ALS

Location	SNP	Odds Ratio*	Genes in locus	Comment	References
chr 19	rs12608932	1.11	UNC13A	Intronic variant, mode of action still unknown	16,17
chr 17	rs35714695	0.88	SARM1, POLDIP2, TMEM199, MIR4723, SEQBQX, VTN, TNFAIP1, KRT18P55, TMEM97, IFT20	Multiple plausible candidate genes	18
chr 21	rs75087725	1.45	C21orf2	Rare coding variant (1-2% minor allele frequency)	19
chr 3	rs616147	1.10	MOBP	Also associated with PSP	124
chr 14	rs10139154	1.09	SCFD1, G2E3	Novel locus	19
chr 12	rs74654358**	1.21	TBK1	Either tagging multiple rare variants or independent actual functional variant	15,50

Table 2 Legend:

Data are biased towards European populations since there is limited genetic evidence from other populations.

Genome-wide association studies identify a variant on chromosome 9p21 as associated with ALS. This variant is a marker that tags the much rarer *C9orf72* hexanucleotide repeat expansion. It is still unknown whether the other associations here also represent tags for one or more rare variants with a large effect on risk or whether the associations are actually functional themselves and are common variants with a small effect on ALS risk. The *C21orf2* variant has no correlation with other nearby variants and thus appears to be identifying a signal confined to *C21orf2*.

Previously associated variants that were not found to be associated with ALS (uncorrected p value > 0.05) in Van Rheenen et al 2016, supplementary table 15),¹⁹ include single nucleotide variants in or near *FGGY*, *ITPR2*, *SUN3*, *C7orf57*, *DPP6*, *CAMK1G*, *SUSD2*, *18q11.2*, *CYP27A1*, *CENPV*, 8q24.13, and the three previously reported variants in *VEGF*;¹²⁵ rs699947, rs1570360, rs2010963, neither were previously reported variants in *PON1* and *PON2*;¹²⁶ rs662,

rs854560, rs10487132), *ANG*;³⁵ rs11701, or *HFE*;¹⁹ rs1799945. Copy number variation in *SMN1* and *SMN2* ^{127,128} and microsatellite variation in *ELP3* ³⁷ have also been reported to be associated with ALS risk, but still await further independent replication by other groups.

* Odds ratios taken from Van Rheenen et al 2016.¹⁹

** This variant was near genome-wide significance (6.6×10^{-8}) in Van Rheenen et al 2016.¹⁹

References

- 1 Phukan, J., Pender, N. P. & Hardiman, O. Cognitive impairment in amyotrophic lateral sclerosis. *The Lancet. Neurology* **6**, 994-1003, doi:10.1016/S1474-4422(07)70265-X (2007).
- 2 Goldstein, L. H. & Abrahams, S. Changes in cognition and behaviour in amyotrophic lateral sclerosis: nature of impairment and implications for assessment. *The Lancet. Neurology* **12**, 368-380, doi:10.1016/S1474-4422(13)70026-7 (2013).
- 3 Johnston, C. A. *et al.* Amyotrophic lateral sclerosis in an urban setting: a population based study of inner city London. *Journal of neurology* **253**, 1642-1643, doi:10.1007/s00415-006-0195-y (2006).
- 4 Marin, B. *et al.* Variation in worldwide incidence of amyotrophic lateral sclerosis: a meta-analysis. *Int J Epidemiol*, doi:10.1093/ije/dyw061 (2016).
- 5 Al-Chalabi, A. & Hardiman, O. The epidemiology of ALS: a conspiracy of genes, environment and time. *Nature reviews. Neurology* **9**, 617-628, doi:10.1038/nrneurol.2013.203 (2013).
- 6 Manjaly, Z. R. *et al.* The sex ratio in amyotrophic lateral sclerosis: A population based study. *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases* **11**, 439-442, doi:10.3109/17482961003610853 (2010).
- 7 Byrne, S. *et al.* Aggregation of neurologic and neuropsychiatric disease in amyotrophic lateral sclerosis kindreds: a population-based case-control cohort study of familial and sporadic amyotrophic lateral sclerosis. *Annals of neurology* **74**, 699-708, doi:10.1002/ana.23969 (2013).
- 8 Al-Chalabi, A. *et al.* An estimate of amyotrophic lateral sclerosis heritability using twin data. *Journal of neurology, neurosurgery, and psychiatry* **81**, 1324-1326, doi:10.1136/jnnp.2010.207464 (2010).
- 9 Wingo, T. S., Cutler, D. J., Yarab, N., Kelly, C. M. & Glass, J. D. The heritability of amyotrophic lateral sclerosis in a clinically ascertained United States research registry. *PloS one* **6**, e27985, doi:10.1371/journal.pone.0027985 (2011).
- 10 McLaughlin, R. L., Vajda, A. & Hardiman, O. Heritability of Amyotrophic Lateral Sclerosis: Insights From Disparate Numbers. *JAMA neurology* **72**, 857-858, doi:10.1001/jamaneurol.2014.4049 (2015).
- 11 Byrne, S., Elamin, M., Bede, P. & Hardiman, O. Absence of consensus in diagnostic criteria for familial neurodegenerative diseases. *Journal of neurology, neurosurgery, and psychiatry* **83**, 365-367, doi:10.1136/jnnp-2011-301530 (2012).
- 12 Al-Chalabi, A. & Lewis, C. M. Modelling the effects of penetrance and family size on rates of sporadic and familial disease. *Human heredity* **71**, 281-288, doi:10.1159/000330167 (2011).
- 13 Andersen, P. M. & Al-Chalabi, A. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nature reviews. Neurology* **7**, 603-615, doi:10.1038/nrneurol.2011.150 (2011).
- 14 Williams, K. L. *et al.* CCNF mutations in amyotrophic lateral sclerosis and frontotemporal dementia. *Nat Commun* **7**, 11253, doi:10.1038/ncomms11253 (2016).
- 15 Freischmidt, A. *et al.* Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia. *Nat Neurosci* **18**, 631-636, doi:10.1038/nn.4000 (2015).

- 16 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. *Nature genetics* **41**, 1083-1087, doi:10.1038/ng.442 (2009).
- 17 Shatunov, A. *et al.* Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *The Lancet. Neurology* **9**, 986-994, doi:10.1016/s1474-4422(10)70197-6 (2010).
- 18 Fogh, I. *et al.* A genome-wide association meta-analysis identifies a novel locus at 17q11.2 associated with sporadic amyotrophic lateral sclerosis. *Human molecular genetics* **23**, 2220-2231, doi:10.1093/hmg/ddt587 (2014).
- 19 van Rheenen, W. *et al.* Genome-wide association analyses identify new risk variants and the genetic architecture of amyotrophic lateral sclerosis. *Nature genetics* **48**, 1043-1048, doi:10.1038/ng.3622 (2016).
- 20 Kenna, K. P. *et al.* NEK1 variants confer susceptibility to amyotrophic lateral sclerosis. *Nature genetics* **48**, 1037-1042, doi:10.1038/ng.3626 (2016).
- 21 Abel, O., Powell, J. F., Andersen, P. M. & Al-Chalabi, A. ALSod: A user-friendly online bioinformatics tool for amyotrophic lateral sclerosis genetics. *Human mutation* **33**, 1345-1351, doi:10.1002/humu.22157 (2012).
- 22 Wroe, R., Wai-Ling Butler, A., Andersen, P. M., Powell, J. F. & Al-Chalabi, A. ALSOD: the Amyotrophic Lateral Sclerosis Online Database. *Amyotrophic lateral sclerosis : official publication of the World Federation of Neurology Research Group on Motor Neuron Diseases* **9**, 249-250, doi:10.1080/17482960802146106 (2008).
- 23 Amberger, J. S., Bocchini, C. A., Schiettecatte, F., Scott, A. F. & Hamosh, A. OMIM.org: Online Mendelian Inheritance in Man (OMIM(R)), an online catalog of human genes and genetic disorders. *Nucleic acids research* **43**, D789-798, doi:10.1093/nar/gku1205 (2015).
- 24 Visscher, P. M., Brown, M. A., McCarthy, M. I. & Yang, J. Five years of GWAS discovery. *American journal of human genetics* **90**, 7-24, doi:10.1016/j.ajhg.2011.11.029 (2012).
- 25 Morita, M. *et al.* A locus on chromosome 9p confers susceptibility to ALS and frontotemporal dementia. *Neurology* **66**, 839-844, doi:10.1212/01.wnl.0000200048.53766.b4 (2006).
- 26 Vance, C. *et al.* Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2-21.3. *Brain : a journal of neurology* **129**, 868-876, doi:10.1093/brain/awl030 (2006).
- 27 Hsiung, G. Y. *et al.* Clinical and pathological features of familial frontotemporal dementia caused by C9ORF72 mutation on chromosome 9p. *Brain : a journal of neurology* **135**, 709-722, doi:10.1093/brain/awr354 (2012).
- 28 Pearson, J. P. *et al.* Familial frontotemporal dementia with amyotrophic lateral sclerosis and a shared haplotype on chromosome 9p. *Journal of neurology* **258**, 647-655, doi:10.1007/s00415-010-5815-x (2011).
- 29 Boxer, A. L. *et al.* Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. *Journal of neurology, neurosurgery, and psychiatry* **82**, 196-203, doi:10.1136/jnnp.2009.204081 (2011).
- 30 Gijssels, I. *et al.* Identification of 2 Loci at chromosomes 9 and 14 in a multiplex family with frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Arch Neurol* **67**, 606-616, doi:10.1001/archneurol.2010.82 (2010).
- 31 Valdmanis, P. N. *et al.* Three families with amyotrophic lateral sclerosis and frontotemporal dementia with evidence of linkage to chromosome 9p. *Arch Neurol* **64**, 240-245, doi:10.1001/archneur.64.2.240 (2007).

- 32 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, doi:10.1016/j.neuron.2011.09.011 (2011).
- 33 Renton, A. E. *et al.* A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron* **72**, 257-268, doi:10.1016/j.neuron.2011.09.010 (2011).
- 34 Greenway, M. J. *et al.* A novel candidate region for ALS on chromosome 14q11.2. *Neurology* **63**, 1936-1938 (2004).
- 35 Greenway, M. J. *et al.* ANG mutations segregate with familial and 'sporadic' amyotrophic lateral sclerosis. *Nature genetics* **38**, 411-413, doi:10.1038/ng1742 (2006).
- 36 Elden, A. C. *et al.* Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS. *Nature* **466**, 1069-1075, doi:10.1038/nature09320 (2010).
- 37 Simpson, C. L. *et al.* Variants of the elongator protein 3 (ELP3) gene are associated with motor neuron degeneration. *Human molecular genetics* **18**, 472-481, doi:10.1093/hmg/ddn375 (2009).
- 38 Blauw, H. M. *et al.* A large genome scan for rare CNVs in amyotrophic lateral sclerosis. *Human molecular genetics* **19**, 4091-4099, doi:10.1093/hmg/ddq323 (2010).
- 39 Al-Chalabi, A. *et al.* Deletions of the heavy neurofilament subunit tail in amyotrophic lateral sclerosis. *Human molecular genetics* **8**, 157-164 (1999).
- 40 Schizophrenia Working Group of the Psychiatric Genomics, C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature* **511**, 421-427, doi:10.1038/nature13595 (2014).
- 41 Lek, M. *et al.* Analysis of protein-coding genetic variation in 60,706 humans. *Nature* **536**, 285-291, doi:10.1038/nature19057 (2016).
- 42 Appel, S. H. & Rowland, L. P. Amyotrophic lateral sclerosis, frontotemporal lobar dementia, and p62: a functional convergence? *Neurology* **79**, 1526-1527, doi:10.1212/WNL.0b013e31826e26ec (2012).
- 43 Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**, 59-62, doi:10.1038/362059a0 (1993).
- 44 Deng, H. X. *et al.* Amyotrophic lateral sclerosis and structural defects in Cu,Zn superoxide dismutase. *Science (New York, N.Y.)* **261**, 1047-1051 (1993).
- 45 Shaw, C. E. *et al.* Mutations in all five exons of SOD-1 may cause ALS. *Annals of neurology* **43**, 390-394, doi:10.1002/ana.410430319 (1998).
- 46 Ionita-Laza, I., McCallum, K., Xu, B. & Buxbaum, J. D. A spectral approach integrating functional genomic annotations for coding and noncoding variants. *Nature genetics* **48**, 214-220, doi:10.1038/ng.3477 (2016).
- 47 Rao, S. S. *et al.* A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**, 1665-1680, doi:10.1016/j.cell.2014.11.021 (2014).
- 48 Carvunis, A. R. *et al.* Evidence for a common evolutionary rate in metazoan transcriptional networks. *Elife* **4**, doi:10.7554/eLife.11615 (2015).
- 49 Campbell, I. M., Shaw, C. A., Stankiewicz, P. & Lupski, J. R. Somatic mosaicism: implications for disease and transmission genetics. *Trends Genet* **31**, 382-392, doi:10.1016/j.tig.2015.03.013 (2015).
- 50 Cirulli, E. T. *et al.* Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways. *Science (New York, N.Y.)* **347**, 1436-1441, doi:10.1126/science.aaa3650 (2015).

- 51 Glass, C. K., Saijo, K., Winner, B., Marchetto, M. C. & Gage, F. H. Mechanisms underlying inflammation in neurodegeneration. *Cell* **140**, 918-934, doi:10.1016/j.cell.2010.02.016 (2010).
- 52 Fang, X. *et al.* The NEK1 interactor, C21ORF2, is required for efficient DNA damage repair. *Acta Biochim Biophys Sin (Shanghai)* **47**, 834-841, doi:10.1093/abbs/gmv076 (2015).
- 53 Lai, C. K. *et al.* Functional characterization of putative cilia genes by high-content analysis. *Mol Biol Cell* **22**, 1104-1119, doi:10.1091/mbc.E10-07-0596 (2011).
- 54 Ma, X., Peterson, R. & Turnbull, J. Adenylyl cyclase type 3, a marker of primary cilia, is reduced in primary cell culture and in lumbar spinal cord in situ in G93A SOD1 mice. *BMC Neurosci* **12**, 71, doi:10.1186/1471-2202-12-71 (2011).
- 55 Smith, B. N. *et al.* Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS. *Neuron* **84**, 324-331, doi:10.1016/j.neuron.2014.09.027 (2014).
- 56 Puls, I. *et al.* Mutant dynactin in motor neuron disease. *Nature genetics* **33**, 455-456, doi:10.1038/ng1123 (2003).
- 57 Hand, C. K. *et al.* Compound heterozygous D90A and D96N SOD1 mutations in a recessive amyotrophic lateral sclerosis family. *Annals of neurology* **49**, 267-271 (2001).
- 58 van Blitterswijk, M. *et al.* Evidence for an oligogenic basis of amyotrophic lateral sclerosis. *Human molecular genetics* **21**, 3776-3784, doi:10.1093/hmg/dds199 (2012).
- 59 van Blitterswijk, M. *et al.* VAPB and C9orf72 mutations in 1 familial amyotrophic lateral sclerosis patient. *Neurobiology of aging* **33**, 2950 e2951-2954, doi:10.1016/j.neurobiolaging.2012.07.004 (2012).
- 60 Bury, J. J. *et al.* Oligogenic inheritance of optineurin (OPTN) and C9ORF72 mutations in ALS highlights localisation of OPTN in the TDP-43-negative inclusions of C9ORF72-ALS. *Neuropathology : official journal of the Japanese Society of Neuropathology* **36**, 125-134, doi:10.1111/neup.12240 (2016).
- 61 Kramer, N. J. *et al.* Spt4 selectively regulates the expression of C9orf72 sense and antisense mutant transcripts. *Science (New York, N.Y.)* **353**, 708-712, doi:10.1126/science.aaf7791 (2016).
- 62 Diekstra, F. P. *et al.* C9orf72 and UNC13A are shared risk loci for amyotrophic lateral sclerosis and frontotemporal dementia: a genome-wide meta-analysis. *Annals of neurology* **76**, 120-133, doi:10.1002/ana.24198 (2014).
- 63 Chio, A. *et al.* Genetic counselling in ALS: facts, uncertainties and clinical suggestions. *Journal of neurology, neurosurgery, and psychiatry* **85**, 478-485, doi:10.1136/jnnp-2013-305546 (2014).
- 64 Smith, B. N. *et al.* The C9ORF72 expansion mutation is a common cause of ALS+/- FTD in Europe and has a single founder. *European journal of human genetics : EJHG* **21**, 102-108, doi:10.1038/ejhg.2012.98 (2013).
- 65 Al-Chalabi, A. *et al.* Analysis of amyotrophic lateral sclerosis as a multistep process: a population-based modelling study. *The Lancet. Neurology* **13**, 1108-1113, doi:10.1016/s1474-4422(14)70219-4 (2014).
- 66 Agosta, F. *et al.* The El Escorial criteria: strengths and weaknesses. *Amyotrophic lateral sclerosis & frontotemporal degeneration* **16**, 1-7, doi:10.3109/21678421.2014.964258 (2015).
- 67 Escott-Price, V. *et al.* Common polygenic variation enhances risk prediction for Alzheimer's disease. *Brain : a journal of neurology* **138**, 3673-3684, doi:10.1093/brain/awv268 (2015).

- 68 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. *The Lancet. Neurology* **12**, 435-442, doi:10.1016/S1474-4422(13)70061-9 (2013).
- 69 Atanasio, A. *et al.* C9orf72 ablation causes immune dysregulation characterized by leukocyte expansion, autoantibody production, and glomerulonephropathy in mice. *Sci Rep* **6**, 23204, doi:10.1038/srep23204 (2016).
- 70 Lopez, E. S. *et al.* Full ablation of C9orf72 in mice causes immune system-related pathology and neoplastic events but no motor neuron defects. *Acta neuropathologica* (2016).
- 71 Doudna, J. A. & Charpentier, E. Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science (New York, N.Y.)* **346**, 1258096, doi:10.1126/science.1258096 (2014).
- 72 Heidenreich, M. & Zhang, F. Applications of CRISPR-Cas systems in neuroscience. *Nat Rev Neurosci* **17**, 36-44, doi:10.1038/nrn.2015.2 (2016).
- 73 Do, R., Kathiresan, S. & Abecasis, G. R. Exome sequencing and complex disease: practical aspects of rare variant association studies. *Human molecular genetics* **21**, R1-9, doi:10.1093/hmg/dds387 (2012).
- 74 Gitcho, M. A. *et al.* TDP-43 A315T mutation in familial motor neuron disease. *Annals of neurology* **63**, 535-538, doi:10.1002/ana.21344 (2008).
- 75 Kabashi, E. *et al.* TARDBP mutations in individuals with sporadic and familial amyotrophic lateral sclerosis. *Nature genetics* **40**, 572-574, doi:10.1038/ng.132 (2008).
- 76 Sreedharan, J. *et al.* TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis. *Science (New York, N.Y.)* **319**, 1668-1672, doi:10.1126/science.1154584 (2008).
- 77 Munch, C. *et al.* Point mutations of the p150 subunit of dynactin (DCTN1) gene in ALS. *Neurology* **63**, 724-726 (2004).
- 78 Al-Chalabi, A. *et al.* Variants in the ALS2 gene are not associated with sporadic amyotrophic lateral sclerosis. *Neurogenetics* **4**, 221-222, doi:10.1007/s10048-003-0152-1 (2003).
- 79 Hadano, S. *et al.* A gene encoding a putative GTPase regulator is mutated in familial amyotrophic lateral sclerosis 2. *Nature genetics* **29**, 166-173, doi:10.1038/ng1001-166 (2001).
- 80 Takahashi, Y. *et al.* ERBB4 mutations that disrupt the neuregulin-ErbB4 pathway cause amyotrophic lateral sclerosis type 19. *American journal of human genetics* **93**, 900-905, doi:10.1016/j.ajhg.2013.09.008 (2013).
- 81 Parkinson, N. *et al.* ALS phenotypes with mutations in CHMP2B (charged multivesicular body protein 2B). *Neurology* **67**, 1074-1077, doi:10.1212/01.wnl.0000231510.89311.8b (2006).
- 82 Brenner, D. *et al.* NEK1 mutations in familial amyotrophic lateral sclerosis. *Brain : a journal of neurology* **139**, e28, doi:10.1093/brain/aww033 (2016).
- 83 Johnson, J. O. *et al.* Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat Neurosci* **17**, 664-666, doi:10.1038/nn.3688 (2014).
- 84 Muller, T. J. *et al.* Phenotype of matrin-3-related distal myopathy in 16 German patients. *Annals of neurology* **76**, 669-680, doi:10.1002/ana.24255 (2014).
- 85 Fecto, F. *et al.* SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis. *Arch Neurol* **68**, 1440-1446, doi:10.1001/archneurol.2011.250 (2011).

- 86 Hirano, M. *et al.* Mutations in the gene encoding p62 in Japanese patients with amyotrophic lateral sclerosis. *Neurology* **80**, 458-463, doi:10.1212/WNL.0b013e31827f0fe5 (2013).
- 87 Le Ber, I. *et al.* SQSTM1 mutations in French patients with frontotemporal dementia or frontotemporal dementia with amyotrophic lateral sclerosis. *JAMA neurology* **70**, 1403-1410, doi:10.1001/jamaneurol.2013.3849 (2013).
- 88 Rubino, E. *et al.* SQSTM1 mutations in frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Neurology* **79**, 1556-1562, doi:10.1212/WNL.0b013e31826e25df (2012).
- 89 Chow, C. Y. *et al.* Deleterious variants of FIG4, a phosphoinositide phosphatase, in patients with ALS. *American journal of human genetics* **84**, 85-88, doi:10.1016/j.ajhg.2008.12.010 (2009).
- 90 Al-Saif, A., Al-Mohanna, F. & Bohlega, S. A mutation in sigma-1 receptor causes juvenile amyotrophic lateral sclerosis. *Annals of neurology* **70**, 913-919, doi:10.1002/ana.22534 (2011).
- 91 Johnson, J. O. *et al.* Exome sequencing reveals VCP mutations as a cause of familial ALS. *Neuron* **68**, 857-864, doi:10.1016/j.neuron.2010.11.036 (2010).
- 92 Kaneb, H. M. *et al.* Deleterious mutations in the essential mRNA metabolism factor, hGle1, in amyotrophic lateral sclerosis. *Human molecular genetics* **24**, 1363-1373, doi:10.1093/hmg/ddu545 (2015).
- 93 Blair, I. P. *et al.* A gene for autosomal dominant juvenile amyotrophic lateral sclerosis (ALS4) localizes to a 500-kb interval on chromosome 9q34. *Neurogenetics* **3**, 1-6 (2000).
- 94 Chance, P. F. *et al.* Linkage of the gene for an autosomal dominant form of juvenile amyotrophic lateral sclerosis to chromosome 9q34. *American journal of human genetics* **62**, 633-640, doi:10.1086/301769 (1998).
- 95 Chen, Y. Z. *et al.* DNA/RNA helicase gene mutations in a form of juvenile amyotrophic lateral sclerosis (ALS4). *American journal of human genetics* **74**, 1128-1135, doi:10.1086/421054 (2004).
- 96 De Jonghe, P. *et al.* Autosomal dominant juvenile amyotrophic lateral sclerosis and distal hereditary motor neuronopathy with pyramidal tract signs: synonyms for the same disorder? *Brain : a journal of neurology* **125**, 1320-1325 (2002).
- 97 Maruyama, H. *et al.* Mutations of optineurin in amyotrophic lateral sclerosis. *Nature* **465**, 223-226, doi:10.1038/nature08971 (2010).
- 98 Pottier, C. *et al.* Whole-genome sequencing reveals important role for TBK1 and OPTN mutations in frontotemporal lobar degeneration without motor neuron disease. *Acta neuropathologica* **130**, 77-92, doi:10.1007/s00401-015-1436-x (2015).
- 99 Gros-Louis, F. *et al.* A frameshift deletion in peripherin gene associated with amyotrophic lateral sclerosis. *J Biol Chem* **279**, 45951-45956, doi:10.1074/jbc.M408139200 (2004).
- 100 Kim, H. J. *et al.* Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* **495**, 467-473, doi:10.1038/nature11922 (2013).
- 101 Daoud, H. *et al.* Association of long ATXN2 CAG repeat sizes with increased risk of amyotrophic lateral sclerosis. *Arch Neurol* **68**, 739-742, doi:10.1001/archneurol.2011.111 (2011).
- 102 Van Damme, P. *et al.* Expanded ATXN2 CAG repeat size in ALS identifies genetic overlap between ALS and SCA2. *Neurology* **76**, 2066-2072, doi:10.1212/WNL.0b013e31821f445b (2011).

- 103 Millecamps, S. *et al.* SOD1, ANG, VAPB, TARDBP, and FUS mutations in familial amyotrophic lateral sclerosis: genotype-phenotype correlations. *Journal of medical genetics* **47**, 554-560, doi:10.1136/jmg.2010.077180 (2010).
- 104 van Es, M. A. *et al.* A case of ALS-FTD in a large FALS pedigree with a K17I ANG mutation. *Neurology* **72**, 287-288, doi:10.1212/01.wnl.0000339487.84908.00 (2009).
- 105 Wu, D. *et al.* Angiogenin loss-of-function mutations in amyotrophic lateral sclerosis. *Annals of neurology* **62**, 609-617, doi:10.1002/ana.21221 (2007).
- 106 Hentati, A. *et al.* Linkage of a commoner form of recessive amyotrophic lateral sclerosis to chromosome 15q15-q22 markers. *Neurogenetics* **2**, 55-60 (1998).
- 107 Orlacchio, A. *et al.* SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis. *Brain : a journal of neurology* **133**, 591-598, doi:10.1093/brain/awp325 (2010).
- 108 Hubers, A. *et al.* De novo FUS mutations are the most frequent genetic cause in early-onset German ALS patients. *Neurobiology of aging* **36**, 3117 e3111-3116, doi:10.1016/j.neurobiolaging.2015.08.005 (2015).
- 109 Kwiatkowski, T. J., Jr. *et al.* Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis. *Science (New York, N.Y.)* **323**, 1205-1208, doi:10.1126/science.1166066 (2009).
- 110 Vance, C. *et al.* Mutations in FUS, an RNA processing protein, cause familial amyotrophic lateral sclerosis type 6. *Science (New York, N.Y.)* **323**, 1208-1211, doi:10.1126/science.1165942 (2009).
- 111 Wu, C. H. *et al.* Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis. *Nature* **488**, 499-503, doi:10.1038/nature11280 (2012).
- 112 Hand, C. K. *et al.* A novel locus for familial amyotrophic lateral sclerosis, on chromosome 18q. *American journal of human genetics* **70**, 251-256 (2002).
- 113 Sapp, P. C. *et al.* Identification of two novel loci for dominantly inherited familial amyotrophic lateral sclerosis. *American journal of human genetics* **73**, 397-403, doi:10.1086/377158 (2003).
- 114 Nishimura, A. L. *et al.* A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. *American journal of human genetics* **75**, 822-831, doi:10.1086/425287 (2004).
- 115 Alexander, M. D. *et al.* "True" sporadic ALS associated with a novel SOD-1 mutation. *Annals of neurology* **52**, 680-683, doi:10.1002/ana.10369 (2002).
- 116 Andersen, P. M. *et al.* Amyotrophic lateral sclerosis associated with homozygosity for an Asp90Ala mutation in CuZn-superoxide dismutase. *Nature genetics* **10**, 61-66, doi:10.1038/ng0595-61 (1995).
- 117 Bannwarth, S. *et al.* A mitochondrial origin for frontotemporal dementia and amyotrophic lateral sclerosis through CHCHD10 involvement. *Brain : a journal of neurology* **137**, 2329-2345, doi:10.1093/brain/awu138 (2014).
- 118 Figlewicz, D. A. *et al.* Variants of the heavy neurofilament subunit are associated with the development of amyotrophic lateral sclerosis. *Human molecular genetics* **3**, 1757-1761 (1994).
- 119 Rooke, K., Figlewicz, D. A., Han, F. Y. & Rouleau, G. A. Analysis of the KSP repeat of the neurofilament heavy subunit in familial amyotrophic lateral sclerosis. *Neurology* **46**, 789-790 (1996).
- 120 Tomkins, J. *et al.* Novel insertion in the KSP region of the neurofilament heavy gene in amyotrophic lateral sclerosis (ALS). *Neuroreport* **9**, 3967-3970 (1998).
- 121 Vechio, J. D., Bruijn, L. I., Xu, Z., Brown, R. H., Jr. & Cleveland, D. W. Sequence variants in human neurofilament proteins: absence of linkage to familial amyotrophic

- lateral sclerosis. *Annals of neurology* **40**, 603-610, doi:10.1002/ana.410400410 (1996).
- 122 Deng, H. X. *et al.* Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia. *Nature* **477**, 211-215, doi:10.1038/nature10353 (2011).
- 123 Fahed, A. C. *et al.* UBQLN2 mutation causing heterogeneous X-linked dominant neurodegeneration. *Annals of neurology* **75**, 793-798, doi:10.1002/ana.24164 (2014).
- 124 Hoglinger, G. U. *et al.* Identification of common variants influencing risk of the tauopathy progressive supranuclear palsy. *Nature genetics* **43**, 699-705, doi:10.1038/ng.859 (2011).
- 125 Lambrechts, D. *et al.* Meta-analysis of vascular endothelial growth factor variations in amyotrophic lateral sclerosis: increased susceptibility in male carriers of the -2578AA genotype. *Journal of medical genetics* **46**, 840-846, doi:10.1136/jmg.2008.058222 (2009).
- 126 Wills, A. M. *et al.* A large-scale international meta-analysis of paraoxonase gene polymorphisms in sporadic ALS. *Neurology* **73**, 16-24, doi:10.1212/WNL.0b013e3181a18674 (2009).
- 127 Blauw, H. M. *et al.* SMN1 gene duplications are associated with sporadic ALS. *Neurology* **78**, 776-780, doi:10.1212/WNL.0b013e318249f697 (2012).
- 128 Corcia, P. *et al.* SMN1 gene, but not SMN2, is a risk factor for sporadic ALS. *Neurology* **67**, 1147-1150, doi:10.1212/01.wnl.0000233830.85206.1e (2006).