



Stratified gene expression analysis identifies major amyotrophic lateral sclerosis genes

Citation

Jones, Ashley R., Claire Troakes, Andrew King, Vibhu Sahni, Simone De Jong, Koen Bossers, Efterpi Papouli, et al. 2015. "Stratified Gene Expression Analysis Identifies Major Amyotrophic Lateral Sclerosis Genes." *Neurobiology of Aging* 36 (5) (May): 2006.e1–2006.e9.

Published Version

doi:10.1016/j.neurobiolaging.2015.02.017

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:17223814>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Open Access Policy Articles, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#OAP>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

Stratified gene expression analysis identifies major amyotrophic lateral sclerosis genes

Ashley R Jones PhD¹, Claire Troakes PhD², Andrew King MD PhD², Vibhu Sahni PhD³, Simone De Jong PhD⁴, Koen Bossers PhD⁵, Efterpi Papouli PhD^{6,7}, Muddassar Mirza MSc⁶, Safa Al-Sarraj PhD FRCPATH², Christopher E Shaw PhD FRCP¹, Pamela J Shaw PhD FRCP⁸, Janine Kirby PhD⁸, Jan Veldink MD PhD⁹, Jeffrey D Macklis MD DHST³, John F Powell PhD¹, Ammar Al-Chalabi PhD FRCP¹

1. King's College London, Institute of Psychiatry, Psychology and Neuroscience, Department of Basic and Clinical Neuroscience, London SE5 8AF, UK
2. MRC London Neurodegenerative Diseases Brain Bank, King's College London, London SE21 8EA, UK.
3. Department of Stem Cell and Regenerative Biology, Center for Brain Science, and Harvard Stem Cell Institute, Harvard University, Cambridge, MA 02138, USA.
4. MRC Social Genetic and Developmental Psychiatry Research Centre, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK.
5. Synaptic plasticity and Behavior group, Netherlands Institute for Neuroscience, Amsterdam 1105, The Netherlands.
6. Biomedical Research Centre, King's College London, 7th Floor Tower Wing, Guy's Hospital, St Thomas Street, London SE1 9RT, UK.
7. Cambridge Epigenetix Ltd, Jonas Webb Building, Babraham Research Campus, CB22 3AT, UK.
8. Sheffield Institute for Translational Neuroscience (SITraN), University of Sheffield, 385A Glossop Road, Sheffield S10 2HQ, UK.
9. Department of Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht 3584, The Netherlands.

Correspondence to:

Ammar Al-Chalabi, Professor of Neurology and Complex Disease Genetics, King's College London
London SE5 8AF, UK, ammar.al-chalabi@kcl.ac.uk +44 20 7848 5187 (t), +44 20 7848 5190 (f)

Abstract

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease of motor neurons resulting in progressive paralysis. Gene expression studies of ALS only rarely identify the same gene pathways as gene association studies. We hypothesised that analysing tissues by matching on degree of disease severity would identify different patterns of gene expression than a traditional case-control comparison. We analysed gene expression changes in four post-mortem CNS regions, stratified by severity of motor neuron loss. An overall case (n = 6) control (n = 3) comparison identified known ALS gene, *SOX5* as showing differential expression (log₂ fold-change = 0.09, p = 5.5x10⁻⁵). Analyses stratified by disease severity identified expression changes in *C9orf72* (p = 2.77x10⁻³), *MATR3* (p = 3.46x10⁻³) and *VEGFA* (p = 8.21x10⁻⁴), both implicated in ALS through genetic studies, as well as changes in other genes in pathways involving RNA processing and immune response. These findings suggest analysis of gene expression stratified by disease severity can identify major ALS genes and may be more efficient than traditional case-control comparison.

Keywords: Gene Expression, Amyotrophic Lateral Sclerosis, Spinal Cord, Disease Spread, C9ORF72, RNA processing

1. Introduction

Amyotrophic Lateral Sclerosis (ALS) is a neurodegenerative disease in which there is progressive loss of upper and lower motor neurons, typically leading to diaphragmatic respiratory failure and death within 3 to 5 years. Risk mutations and other genetic variants have now been identified in a substantial proportion of patients, and more than 100 genes have been implicated (Abel, et al., 2012, Andersen and Al-Chalabi, 2011). Common mechanisms such as accumulation of phosphorylated TAR DNA-binding protein 43 (TDP-

43) proteins are being identified, despite substantial genetic, cellular and clinical heterogeneity (Ravits, et al., 2013).

ALS is an adult-onset condition, even though the major predisposing factor maybe present from early life (Al-Chalabi, et al., 2014). This is clearly seen in familial ALS where the genetic lesion is present from conception (Al-Chalabi and Hardiman, 2013). Furthermore, despite the lesion being present in every cell, ALS onset seems to occur in one neurological segment, spreading to neighbouring regions within a predictable portion of time (Roche, et al., 2012). The pattern of clinical spread is also predictable when using neuroanatomy as a framework (Ravits, et al., 2007). Neuropathologically, these findings suggest that ALS begins focally and spreads diffusely throughout the corticospinal, bulbar and spinal motor network.

If the contiguous spread model of ALS is true, then anatomical regions close to the site of onset would be at a more advanced disease stage than topographically distant regions, and should show a different gene expression profile from those at an earlier disease stage regardless of anatomical segment. We therefore analysed gene expression in cases and controls stratified by anatomical segment, and compared these results with the analysis stratified by severity of disease in affected tissue as defined by both distance from site of onset and neuropathological examination.

2. Materials & methods

2.1 Setting and Patients

Brain and spinal cord samples were from the Medical Research Council's London Brain Bank for Neurodegenerative Disease based at the Institute of Psychiatry, King's College London, or Brains for Dementia Research at King's College London (see Table 1 for demographic and disease information). All

Spinal Cord Gene Expression in ALS

patients had sporadic and limb-onset ALS. From each patient, four nervous system segments were isolated and analysed: Medulla, Cervical, Thoracic and Lumbar.

2.2 Tissue repository and RNA and DNA isolation

Tissue was either flash frozen post-mortem and stored at -80°C or formalin-fixed and paraffin embedded (FFPE) blocks. Frozen tissue blocks of 20mg were taken from available medullary, cervical, thoracic and lumbar regions. RNA isolation was performed by submersing the 20mg tissue block in 900µl QIAzol lysis reagent within lysing matrix D (MP Biomedicals, Santa Ana, CA, USA). The tissue in the lysing matrix was homogenised in FastPrep 24 (MP Biomedicals, Santa Ana, CA, USA) for 30 seconds at 4 metres per second. RNA was isolated using RNeasy Universal Kit (Qiagen, Valencia, CA, US) using the manufacturer's protocol. RNA was stored at -80°C in RNase-free water.

DNA was isolated by submersing the 25mg frozen tissue block in 80µl PBS and homogenising the tissue using a rotor-stator homogeniser, followed by extraction using the QIAamp DNA Mini and Blood Mini Kit (Qiagen, Valencia, CA, US). Proteinase K was added to the homogenate to deactivate protein activity.

2.3 RNA quantification and quality control

RNA quantification was estimated using a Life Technologies Qubit 2.0 fluorometer and kit reagents (Carlsband, CA, USA). DNA and RNA quality was examined via 260/280 absorbance ratios using a Thermo Scientific Nanodrop (Waltham, MA, USA) and DNA quality and RNA integrity (RIN) using an Agilent 2100 Bioanalyser (Santa Clara, CA, USA).

2.4 Histological Analyses

Haematoxylin and eosin (H&E) staining of FFPE blocks was used to estimate the extent of loss of motor neurons in each spinal and bulbar/medulla sample for all cases. Using a semi-quantitative score each sample was categorised into one of four grades of severity (see Table S1):

Mild loss of motor neurons: Few neurons seen with angulation and /or chromatolysis.

Moderate loss of motor neurons: Moderate numbers of angulated neurons with chromatolysis and occasional atrophic neurons.

Moderate-severe loss of motor neurons: Angulated and atrophic neurons obvious, with 20% of surviving neurons with normal features. Areas indicating neuronal loss obvious.

Severe loss of motor neurons: Very few surviving neurons. Fewer than 20% of those surviving having a normal appearance on H&E, with some evidence of gliosis and/or neuronophagia.

2.5 Whole-Genome Gene Expression using Illumina DASL HT Assays

We used Illumina Human Whole-Genome DASL HT Assay with UDG kit, containing protocol, reagents and BeadChips for the expression analysis, on the Illumina BeadArray platform (San Diego, CA, USA)

2.6 Probe and sample quality control

Illumina GenomeStudio 2011.1 was used to examine the quality of the RNA expression data. As sample quality control, background subtraction was performed and outliers were identified by examining signal averages of control probes by case, and removed accordingly. As probe quality control, to assess the expression data quality and number of probe for each case, the housekeeping probe average signal was compared to background noise. The 95th percentile was compared to background noise using as a signal-to-noise ratio to assess quality of expression, and to assess the strength of probe signals. An overall

Spinal Cord Gene Expression in ALS

average signal was examined using a box-plot; where probes had an average signal $> 64,000$ were deleted as recommended by Illumina.

For sample quality control, cluster analyses and related dendrograms were used to help to confirm biological replicates within individuals and identify any significant outliers. This analysis used the metric $1 - r$, r being a correlation coefficient of probe expression, for all cases and controls. A scatter plot was performed to examine signal intensities across two samples at a time, with exclusion criteria of $r < 0.99$. Only probes with a detection p-values < 0.05 were selected.

Bioconductor R Package Lumi 2.14 (Du, et al., 2008) was utilised for additional quality control steps, which included variance stabilising transformation and quantile normalisation, and obtaining a quality control estimate. Density, cumulative distribution function, and sample hierarchical clustering plots were drawn to highlight possible abnormalities in expression data.

2.7 Statistical analyses

Gene expression data was analysed using three approaches: first, by comparing expression from all tissue samples between cases and controls irrespective of their location (a typical expression study design); second, by comparing expression in each of the four segments individually between cases and controls; and third, by grouping spinal tissue by grade of severity, comparing gene expression in each severity grade with that in controls, and analysing changes in expression across the four grades of severity.

Differential expression analyses were run using Bioconductor R package Limma 2.14. This incorporates single-channel linear modelling (lmfit), comparing log-intensities of gene expression using make.contrasts

Spinal Cord Gene Expression in ALS

which also calculates a t-statistic, and a post-analyses Bayesian modification of the t-statistic and log-odds for differential expression. Hierarchical clustering and heatmaps were created using heatmap.2 from the R package gplots(Warnes, et al., 2009). These tests were used for all comparisons except the analysis in expression across the four grades of severity.

To test whether gene expression change was dose-dependent across different grades of severity we used a repeated-measures genome-wide one-way ANOVA with gene expression level as the dependent variable and four grades of severity, which were ascribed by histological analysis. Using probes with p-values < 0.001 from the ANOVA analysis, we used Bioconductor package Heatmap.2 (Warnes, et al., 2009) and hclust, from which we were able to cluster and display probes in terms of their expression profile throughout progressing severity. Using probes with p-values < 0.01 from the ANOVA analysis, MFuzz(Futschik, 2012), a time-based soft-clustering algorithm for gene expression, was used to identify clusters of genes which share similar changes in expression as a consequence of tissue severity grade.

Since RNA expression changes are likely to occur within networks, statistically significant differentially expressed genes were analysed using DAVID 6.7(Huang da, et al., 2009) and AmiGO, for annotation of the clusters identified by DAVID. To allow DAVID's functional annotation clustering in analyses where there were few statistically significant changes in gene expression, the non-adjusted p-value threshold were lowered to $p < 0.001$, and $p < 0.01$ where necessary.

2.8 Analyses using network tools

Genes of interest were examined using the protein-protein network tool STRING 9.05(Jensen, et al., 2009) and the gene-gene and gene-protein network tool GeneMANIA(Warde-Farley, et al., 2010). Networks showing relevance with previous ALS research are reported.

3. Results

3.1. Patient Characteristics and Quality Control

There were 6 patients and 3 controls. Mean age of onset for the patients was 60.8 (Table 1). Mean post mortem delay was 28.8 hours for cases, and 35 for controls. Mean age at death was 64 years for cases and 64.6 for controls.

One tissue sample (mild affected cervical tissue) from patient ALS_3 was excluded from all analyses, as it was an outlier in the principal components quality control step. Three tissue samples were excluded from patients ALS_1, ALS_2, and ALS_3 as their degree of motor neuron loss could not be established (Table S1). Before quality control there were 29376 probes, representing 20818 genes. After quality control 20040 probes remained, representing 14707 genes.

	Sex	Age at death (yrs)	Post Mortem Delay (hrs)	AOO (yrs)	ALS Type
ALS_1	F	75	38	72.4	Limb onset
ALS_2	F	80	37	65.4	Limb onset
ALS_3	F	65	14	61.3	Limb onset

Spinal Cord Gene Expression in ALS

ALS_4	F	63	25	58.6	Limb onset
ALS_5	M	50	26	46.5	Limb onset
ALS_6	F	51	33	N/A	ALS-FTD
CTRL_1	F	54	31	-	-
CTRL_2	M	89	41	-	-
CTRL_3	F	51	33	-	-

Table 1. Case and controls demographic information, with disease information for cases.

3.2. Case-control analysis

3.2.1. Analysis regardless of anatomical segment or severity

In the first instance, we compared tissue samples from all cases (n = 6, samples n = 20) with all controls (n = 3, samples n = 12) to identify genes showing differential expression between ALS and control samples regardless of anatomical segment, in a typical gene expression study design. After correction for multiple testing, 153 probes showed statistically significant differential expression with an adjusted p < 0.05, representing 143 genes (see Supplementary Table S2). There were 7 functional annotation clusters with an enrichment score > 1.3. The main cluster identified was involved in insulin-like growth factor binding (enrichment score = 1.49; see Supplementary Table S3).

3.2.2. Analysis by anatomical segment

3.2.2.1. Analysis of medullary samples

Comparing cases (n = 5) with controls (n = 12) using medullary samples identified no probe that showed statistically significant differential expression after multiple-testing correction. We selected probes with a

Spinal Cord Gene Expression in ALS

non-adjusted p-value < 0.01; there were 43 probes representing 42 genes (see Supplementary Table S4). Functional annotation clustering using DAVID identified 3 clusters with an enrichment score greater than 1.3. The main cluster identified was involved in the glycoproteins and signalling (enrichment score = 2.51; see Supplementary Table S5).

3.2.2.2. Analysis of Cervical samples

Comparing cases (n = 5) with controls (n = 12) using cervical samples identified no probe that showed statistically significant differential expression after multiple-testing correction. We selected probes that showed differential expression with a non-adjusted p-value < 0.01; there were 69 probes representing 66 genes (see Supplementary Table S6). Functional annotation clustering using DAVID identified 10 clusters with an enrichment score greater than 1.3. The main cluster identified was involved in signal peptides and the extracellular region (enrichment score = 4.2, see Supplementary Table S7).

The gene *Adenosine Deaminase (ADA)*, which differentially expressed in this category, was predicted to interact with ALS gene *Amyotrophic Lateral Sclerosis 2 (juvenile) (ALS2)* and is known to interact with *Dipeptidyl-Peptidase 4 (DPP4)* (Fig. S1). No other cluster had a reasonably high enrichment score or showed relevance in terms of function with known ALS pathology.

3.2.2.3. Analysis of Thoracic samples

Comparing cases (n = 6) with controls (n = 12) using thoracic samples identified no probe that showed statistically significant differential expression after multiple-testing correction. We selected probes that showed differential expression with a non-adjusted p-value < 0.01; there were 565 probes representing 549 genes (see Supplementary Table S8). Functional annotation clustering using DAVID identified 7 clusters

Spinal Cord Gene Expression in ALS

with an enrichment score greater than 1.3. The main cluster identified was involved in osteoblast differentiation (enrichment score = 2.23, see Supplementary Table S9).

As well as up-regulation of *SRY (sex determining region Y)-box 5 (SOX5)* expression, as seen earlier, three other genes previously implicated in ALS were found to differentially express. These were *Zinc Finger Protein 64 Homolog (ZFP64)*, *Vacuolar Protein Sorting 54 Homolog (VPS54)*, and *Disrupted in Schizophrenia 1 (DISC1)*. The gene *Ubiquitin Fusion Degradation 1 Like (yeast) (UFD1L)*, which also differentially expressed, binds with *Valosin-Containing Protein (VCP)* to form a protein complex (see Fig. S2). No other cluster had an enrichment score over 1.3 or showed relevance in terms of function with known ALS pathology.

3.2.2.4. Analysis of Lumbar samples

Comparing cases (n = 4) with controls (n = 12) using lumbar samples identified no probe that showed statistically significant differential expression after multiple-testing correction. We selected probes that showed differential expression with a non-adjusted p-value < 0.01; there were 276 probes representing 267 genes (see Supplementary Table S10). Functional annotation clustering using DAVID identified three functional clusters with an enrichment score greater than 1.3. The most significant cluster was involved in amino acid transmembrane activity (enrichment score = 2.03, see Supplementary Table S11).

The gene *Cytochrome B5 Reductase 1 (CYB5R1)*, which showed differential expression in this cluster, binds with *Ubiquilin 4 (UBQLN4)* which, similarly to *VCP* and *Ubiquilin 2 (UBQLN2)*, regulates proteosomal protein catabolic processes (Fig. S3), and may be a possible candidate for ALS.

Spinal Cord Gene Expression in ALS

For lumbar samples, there were five genes previously implicated in ALS that showed significant differential expression, all of which were down-regulated. These were *Dynactin 1 (DCTN1)*, *Paraoxonase 1 (PON1)*, *Poliovirus Receptor (PVR)*, *Solute Carrier Family 1 (glial high affinity glutamate transporter), Member 2 (SLC1A2)* and *Unc-13 Homolog A (UNC13A)*.

3.3. Analyses of expression by grade of severity

See Table S1 for pathological and clinical information for each case by spinal segment. To control for genes that naturally show differential expression between anatomical segments, probe expression was compared in a 4x4 matrix in control samples, including each of the levels of spinal cord. 149 genes were then excluded from subsequent analyses (Supplementary Table S12) as they showed inter-sample differential expression. Thus the only genes left in analyses were those that might vary by severity rather than nervous system location.

3.3.1. Analysis of Mild severity grade Samples

We compared tissue samples with mild motor neuron loss (n = 4) with control tissue samples (n = 12). After correcting for multiple testing, 54 probes showed statistically significant differential expression. These probes represented 51 genes (see Supplementary Table S13). Functional annotation clustering using DAVID identified 4 clusters with an enrichment score greater than 1.3. This cluster was involved in membrane and glycosylation (see Supplementary Table S14).

3.3.2. Analysis of Moderate severity grade Samples

We compared tissue samples with moderate motor neuron loss (n = 5) with control tissue samples (n = 12). After correcting for multiple testing, 3 probes showed statistically significant differential expression. These

Spinal Cord Gene Expression in ALS

probes represented 3 genes, which were *Leucine-rich Repeat LGI Family, Member 4 (LGI4)*, *Non-Protein Coding RNA 95 (NCRNA00095)*, and *Solute Carrier Family 1 (Glutamate Transporter), Member 7 (SLC1A7)*, of all which down-regulated. Functional annotation clustering using DAVID was not performed due to expected lack of statistical power.

3.3.3. Analysis of Moderate-Severe severity grade Samples

We compared tissue samples with moderate-severe motor neuron loss (n = 7) with control tissue samples (n = 12). After correcting for multiple testing, 15 probes showed statistically significant differential expression. These probes represented 15 genes (see Supplementary Table S15). Functional annotation clustering using DAVID identified four functional clusters with an enrichment score greater than 1.3. This cluster was involved in transmembrane glycosylation (see Supplementary Table S16).

3.3.3.4. Analysis of Severe severity grade Samples

We compared tissue samples with severe motor neuron loss (n = 4) with control tissue samples (n = 12). After correcting for multiple testing, 31 probes showed statistically significant differential expression. These probes represented 27 genes (see Supplementary Table S17). The most significantly expressed gene, *Angiopoietin 2 (ANGPT2)*, was represented three times in this category and interacts with proposed ALS gene *Vascular Endothelial Growth Factor A (VEGFA)*. Functional annotation clustering using DAVID did not identify any cluster with an enrichment >1.3. Clusters with an enrichment <1.3 were involved in ion homeostasis and apoptosis immune response.

3.3.3.5. Analysis of grade-dependent changes in gene expression

Spinal Cord Gene Expression in ALS

The gene *SLC1A7*, showed consistent significant down-regulation throughout each grade of tissue of severity (Fig. 1). *Cholinergic Receptor, Nicotinic, Alpha 1 (Muscle) (CHRNA1)*, showed up-regulation in three grades. Fig. 1 also shows significant differential expression of *ANGPT2* in two severity grades. The *ANGPT2* protein is known to interact with *VEGFA* (Fig. S4). *ANGPT2* is down-regulated towards more severe grades of motor neuron loss, but as we see from the following analysis, *VEGFA* is up-regulated to begin with but later decreases, and shows statistically significant changes between severity grades when modelled in a repeated measures ANOVA.

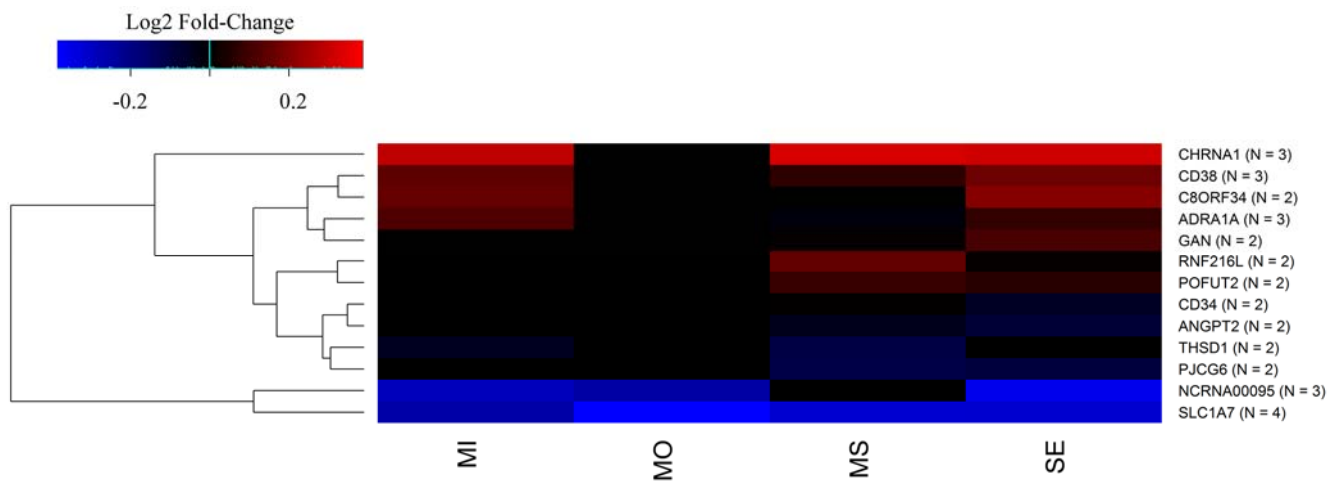


Figure 1. Log Fold-Change by severity grade for genes showing statistical differential expression in more than two grades. MI = mild grade; MO = moderate grade; MS = moderate-severe grade; SE = severe grade. The scale represents log2 fold change, with blue representing down-regulation and red representing up-regulation on a spectrum.

The repeated-measures genome-wide ANOVA of expression changes between each grade revealed 62 genes for which expression statistically significantly changed (Supplementary Table S18). Hierarchical cluster analysis of this cohort revealed two groups; Cluster 1 (n = 37; DAVID functional enrichment score =

Spinal Cord Gene Expression in ALS

1.11) describes a gradual increase in function intracellular protein transport & localisation, and Cluster 2 (n = 12; DAVID functional enrichment score = 1.22) up-down change in leukocyte response. See Fig. 2.

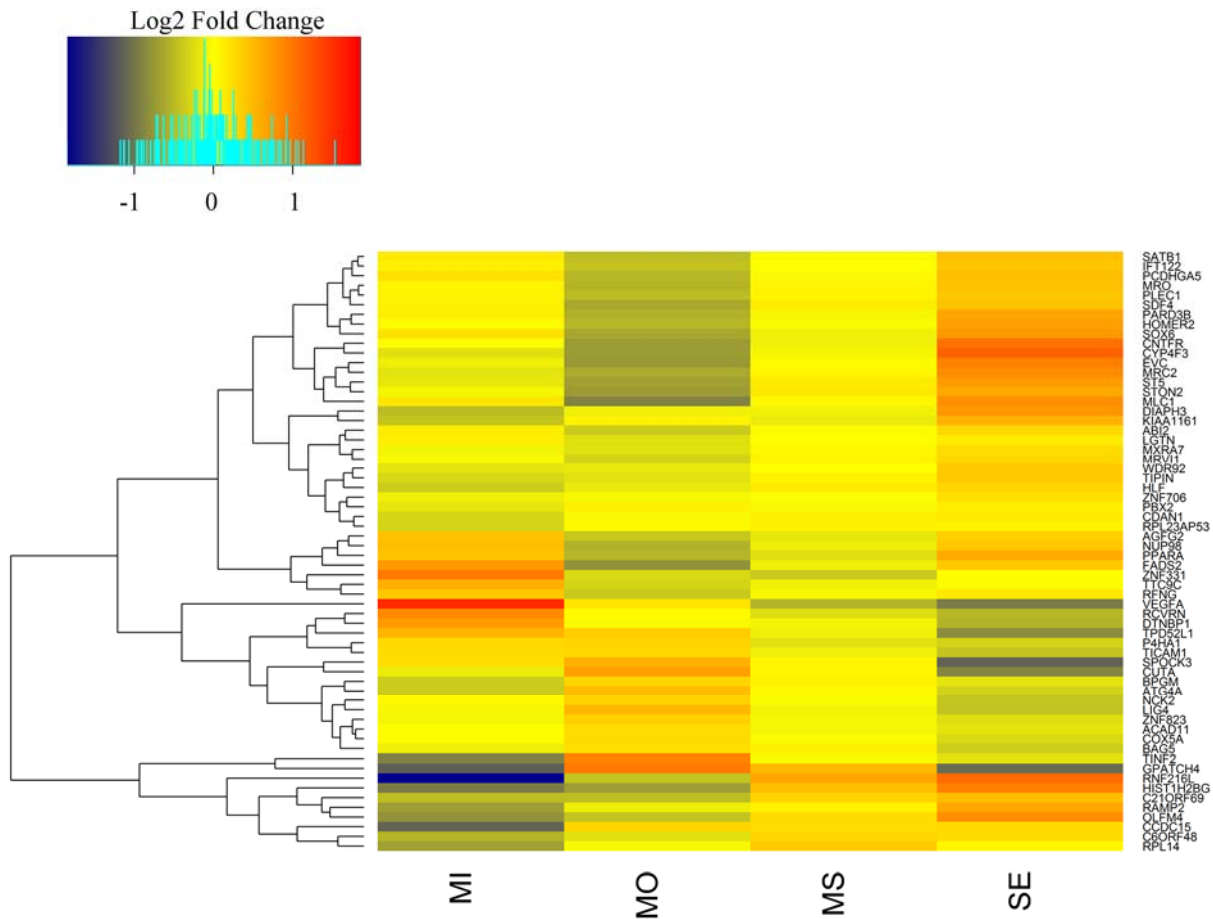


Figure 2. Heatmap of change in gene expression compared to average gene expression from genome-wide one-way ANOVA. Only genes with $p < 0.001$ included. Left y-axis represents cluster dendrogram branches based on similar patterns of expression change across severity grades (x-axis). The right-axis gives the list of genes each branch represents. MI = mild grade; MO = moderate grade; MS = moderate-severe grade; SE = severe grade.

Cluster analysis on the change in expression was performed, using soft-clustering via the R package MFuzz (see Fig. 3). Genes were selected from the previous genome-wide ANOVA with the criteria of a non-

Spinal Cord Gene Expression in ALS

adjusted $p < 0.01$ ($n = 474$) (see Supplementary Table S19). The analysis differentiated five clusters based on their expression pattern across severity grades.

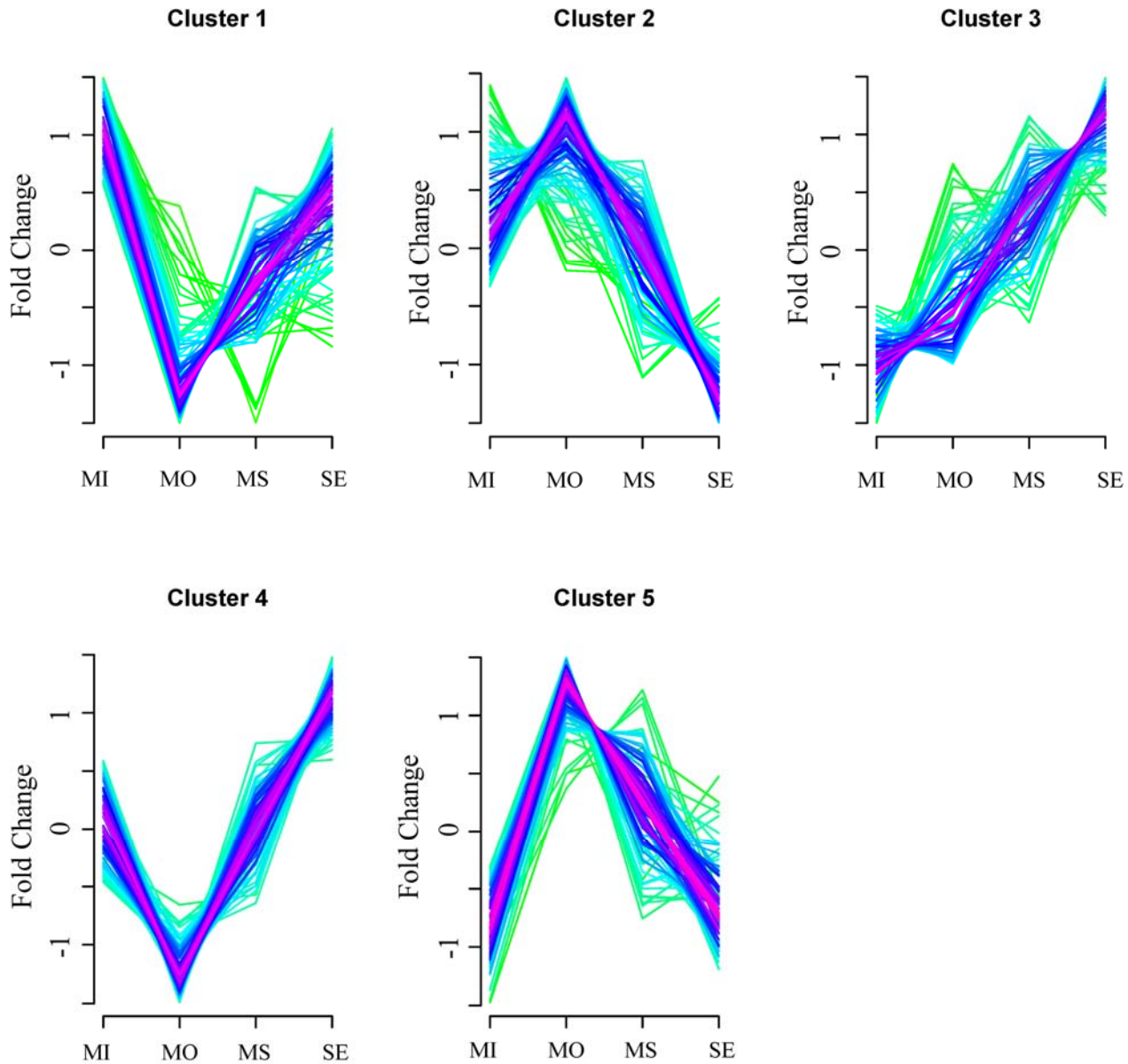


Figure 3. MFuzz clustering by pattern of gene expression. MI = mild grade; MO = moderate grade; MS = moderate-severe grade; SE = severe grade. Colours from purple to green represent categories defined by probe membership values. Purple have membership value towards 1.

Spinal Cord Gene Expression in ALS

The genes of each cluster were analysed for enrichment using DAVID; the following list shows the most statistically enriched categories, with change moving from mild to severe grades of motor neuron loss:

Cluster 1. Down-up change in the regulation of GTPase activity (enrichment score = 1.64)

Cluster 2. Up-down change in the regulation of leukocytes proliferation (enrichment score = 1.52) & regulation of transcription activity (enrichment score = 1.39)

Cluster 3. Linear increase in nucleic methylation (enrichment score = 1.6)

Cluster 4. Down-up change in calcium related cell-adhesion (enrichment score = 1.67)

Cluster 5. Down-up-down change in nucleic lumen function (enrichment score = 1.56)

MFuzz cluster 2 contains two major ALS genes, *Chromosome 9 open reading frame 72 (C9orf72)* ($p = 0.003$) and *Matrin 3 (MATR3)* ($p = 0.003$), and one ALS-related gene *Solute Carrier Family 39 (Metal Ion Transporter), Member 11 (SLC39A11)* ($p = 0.002$).

4. Discussion

We have shown that gene expression analysis in post-mortem ALS tissue, stratified by disease severity as defined both by neuropathology and by distance from onset site, identifies major ALS genes (Table 2) that are not identified when comparing expression between cases and controls at the individual spinal level. Furthermore, functional annotation of genes based on expression change highlighted pathways known to be important in ALS, such as RNA processing and immune response. Direct comparison of expression between tissue-type did not identify these pathways. Although previous gene expression studies have examined changes in ALS spinal cord and bulbar tissue, our strategy has been to identify changes in expression with disease progression, using the hypothesis of disease spread from a focus to allow

Gene	Category	Putative Function	Ref
<i>Differential expression analyses comparing cases with control using all spinal segments (p < 0.05 after multiple testing correction)</i>			
SOX5	N/A	Transcription factor for corticofugal neuron subtype specification	(Daoud, et al., 2011,Lai, et al., 2008)
<i>Differential expression analyses comparing cases with controls by spinal segments (p < 0.01)</i>			
DISC1	Thoracic	Neurite outgrowth and cortical development	(Landers, et al., 2009)
SOX5	Thoracic	Transcription factor for corticofugal neuron subtype development	(Daoud, et al., 2011,Lai, et al., 2008)
VPS54	Thoracic	Retrograde transport of proteins from prevacuoles to late Golgi	(Meisler, et al., 2008)
ZFP64	Thoracic	May be involved in transcriptional regulation	(Schymick, et al., 2007)
DCTN1	Lumbar	Endoplasmic-reticulum-to-Golgi transport, and axonogenesis	(Munch, et al., 2004,Puls, et al., 2003)
PON1	Lumbar	Hydrolyses the toxic metabolites of organophosphorus Insecticides	(Saeed, et al., 2006)
PVR	Lumbar	Transmembrane glycoprotein from the immunoglobulin superfamily	(Saunderson, et al., 2004)
SLC1A2	Lumbar	Transports glutamate from the synaptic extracellular space	(Lin, et al., 1998,Meyer, et al., 1999)
UNC13A	Lumbar	Vesicle maturation during exocytosis	(Chiò, et al., 2009,Daoud, et al., 2010,Shatunov, et al., 2010,van Es, et al., 2009)
<i>Differential expression analyses comparing cases with controls by severity of affected tissue (p < 0.05 after multiple testing correction)</i>			
VEGFA	Mild-mod	Growth factor in angiogenesis	(Lambrechts, et al., 2003,Oosthuyse, et al., 2001)

Spinal Cord Gene Expression in ALS

ANOVA grade-dependent changes in probe expression - ANOVA ($p < 0.001$)

<i>VEGFA</i>	N/A	Growth factor in angiogenesis	(Lambrechts, et al., 2003,Oosthuyse, et al., 2001)
<i>DIAPH3</i>	N/A	Actin remodelling, and regulates cell movement and adhesion	(Daoud, et al., 2011)

MFuzz clustering of grade-dependent changes in probe expression using ANOVA $p < 0.01$

<i>C9orf72</i>	Cluster 2	Possible intracellular membrane trafficking	(DeJesus-Hernandez, et al., 2011,Renton, et al., 2011)
<i>DIAPH3</i>	N/A	Actin remodelling, and regulates cell movement and adhesion	(Daoud, et al., 2011)
<i>HNRNPA1</i>	N/A	Pre-mRNA binding protein involved in metabolism & transport	(Kim, et al., 2013)
<i>MATR3</i>	Cluster 2	May stabilise certain mRNA species or play a role in transcription	(Johnson, et al., 2014)
<i>SLC39A11</i>	Cluster 2	Metal ion transmembrane transporter activity	(Landers, et al., 2009)
<i>VEGFA</i>	N/A	Growth factor in angiogenesis	(Lambrechts, et al., 2003,Oosthuyse, et al., 2001)

Table 2. Genes showing statistical significance that have previously been implicated in ALS.

topographical distance from onset to act as a proxy for disease stage. The pathological grading of nervous system segments followed our expectations under the hypothesis of spread, supporting this approach. These findings suggest that future analyses should take into account local disease stage, and lend support to the hypothesis of disease spread from a focus.

One hundred and forty-three genes showed changes in expression when comparing all tissue samples between cases and controls. The greatest change was seen for *SLC1A7*, showing a log₂ fold change of -0.33. *SLC1A7* was consistently down-regulated across most analyses, including statistically significant expression changes in cervical, thoracic and lumbar tissues, and in all severity grades. As it was consistently down-regulated, we could not see significant changes in expression in grade-dependent analyses. *SLC1A7* has been implicated in ALS previously in a comparative genomic hybridisation study examining genome-wide copy number variation in 71 sporadic ALS cases and 700 non-ALS controls (Shoichet, et al., 2009). *SLC1A7* protects motor neurons from excitotoxicity by transporting extracellular glutamate (Kanai, et al., 2013) through its protein Excitatory Amino-Acid Transporter 5 (EAAT5), and therefore could be involved in ALS pathogenesis through the excitotoxicity hypothesis. EAAT5 shares 36% sequence identity with glia protein Excitatory Amino-Acid Transporter 2 (EAAT2), which is encoded by *SLC1A2* of the same gene family and has almost the exact same function as EAAT5. *SLC1A2* was first implicated in ALS in 1996 (Bristol and Rothstein, 1996).

SOX5 showed significant up-regulation of log₂ fold-change 0.09, and has been previously implicated in ALS through candidate gene sequencing (Daoud, et al., 2011). *SOX5* is a developmental gene involved in early corticospinal motor neuron specification. For discussion on gene ontology enrichment analyses see supplementary section 1.

Spinal Cord Gene Expression in ALS

Comparing probe expression between cases and controls for each separate spinal segment revealed several noteworthy genes, two of which are major ALS genes (*DCTN1* and *UNC13A*), seven genes implicated in ALS previously (*DISC1*, *SOX5*, *VPS54*, *ZFP64*, *PON1*, *PVR* and *SLC1A2*), two candidate genes showing putative interactions with major ALS genes (*ADA* and *UFD1L*), and one candidate showing an interaction with *UBQLN4* and that regulates proteosomal protein catabolic processes (*CYB5R1*). For further discussion on these genes see supplementary section 2, and for gene ontology enrichment analyses section 3.

Analyses comparing tissue samples categorised by their severity-grade against non-ALS control samples revealed *CHRNA1* consistently the most up-regulated gene in all severity grades except moderate-affected tissue. It was the fourth most statistically significant gene when comparing all cases and controls. *CHRNA1* has been found to be up-regulated in previous ALS studies (Bernardini, et al., 2013, Bruneteau, et al., 2013). The gene encodes a subunit of muscle acetylcholine receptors that are involved in binding. For discussion of gene *NCRNA00095* see supplementary section 4, and supplementary section 5 for ontology analyses using severity-grading.

We identified significant changes in *ANGPT2* expression. *ANGPT2* is known to interact with *VEGFA* and their relationship seems bidirectional (Fig. S4). Lack of inhibitory mechanisms by *VEGFA* upon *ANGPT2* induces endothelial cell apoptosis. However interactions between *VEGFA* and *ANGPT2* facilitates endothelial cell migration and proliferation. *ANGPT2* was down-regulated in moderate-severe and severe grades of motor neuron loss. *VEGFA* is up-regulated to begin with but later decreases, showing significant changes in expression across severity grades. Normal expression of *VEGFA* and *ANGPT2* should facilitate healthy inhibition of apoptosis and cell proliferation. If *VEGFA* is over-expressing and is overly inhibitory

Spinal Cord Gene Expression in ALS

towards *ANGPT2* then it could result in failed apoptotic responses. Furthermore, as ALS progresses, down-regulation of either *VEGFA* or *ANGPT2* upon the other could cause improper apoptosis or dysfunctional cell proliferation.

For MFuzz time-based soft-clustering, five clusters were identified. Cluster 2 showed an up-down change in expression as the severity of the tissue worsened. This is the only cluster from the time-based MFuzz analyses to contain ALS genes, two major (*C9orf72* and *MATR3*) and one implicated (*SLC39A11*). Point mutations in *MATR3* have been associated with ALS, and *MATR3* pathology is seen in spinal cord from patients with or without these mutations. *MATR3* encodes for an RNA-binding protein that interacts with the ALS protein TDP-43 (Johnson, et al., 2014).

Currently, *C9orf72* is the most important gene in ALS genetics explaining the largest portion of Caucasian ALS cases (Majounie, et al., 2012, Shatunov, et al., 2010). Differential expression analyses comparing cases and controls have been mixed (Belzil, et al., 2013, Renton, et al., 2011) but most evidence suggests a reduction in *C9orf72* transcripts (Ciura, et al., 2013, DeJesus-Hernandez, et al., 2011, Donnelly, et al., 2013, Lagier-Tourenne, et al., 2013). We recorded an up-down change in *C9orf72* expression as tissue became more severely affected. The expression change was statistically significant and was in the top 0.77% for genes in which expression changed as a function of increasing severity. It followed a common expression pattern with 48 other genes belonging to Cluster 2 from the MFuzz analysis, and it is possible that *C9orf72* aberrant expression may be consequential on some of these genes if they co-express with one another.

Spinal Cord Gene Expression in ALS

If *C9orf72* pathology is due to a toxic-gain-of-function then these results suggest that *C9orf72* pathology is heightened in mild affected (onset) regions, where it is up-regulated, in comparison to its severely affected regions (late stage disease), where it is down-regulated. This may reflect the depletion of motor neurons as the disease worsens or a decrease in the expression of *C9orf72* due to an unknown factor.

A dilemma for gene expression studies is the inference of causality. Functional categories such as angiogenesis and the involvement of genes like *VEGFA* may be a consequence of ALS, rather than its cause. However, many of the genes highlighted in this paper and the functional categories they form, are known to harbour genetic mutations in ALS, which suggests that causality follows the central dogma of genetic mutation to dysfunctional gene expression.

A limitation of this study is the sample size. Post mortem ALS tissue does not routinely include spinal cord, and the ability to extract high quality RNA is affected by multiple factors. Nevertheless, it is encouraging that known ALS genes are identified using this technique. A second limitation of our analyses, which could be problematic for stratified analyses in grade-dependent changes in expression, is the treatment of medullary samples as spinal samples. From our data, this strategy seems reasonable, as functional categories and genes significant in the medullary samples were also significant in non-medullary samples.

A third issue is how to interpret results from tissue samples which have a severe reduction of motor neurons. Notably, there were no highly enriched functional categories for genes showing significant fold-change in severe-grade samples; the most enriched category was involved in disulphide bonds. Yet, there were no samples with complete motor neuron depletion and genes found to differentially express in the moderate-severe grade were also found to be differentially expressed in the severe grade. Interpretation

Spinal Cord Gene Expression in ALS

of findings in this severity grade should however be considered taking into account the scarcity of motor neurons in these samples. Genes found to differentially express in moderate-severe to severe categories are likely due to astrocyte expression.

We were able to identify more genes showing differential expression by stratifying tissue samples by severity of motor neuron loss rather than by tissue type. We have tried to account for differences in expression between nervous system segments by excluding genes showing inter-regional expression changes in controls. Thus, the only remaining difference between segments is disease stage. The ontologically enriched categories derived from the analyses by severity grade, such as glycosylation and transmembrane activity, angiogenesis and immune response, and the finding that many known or proposed ALS genes show expression changes with changes in disease severity, support the evidence that the genes identified in these analyses are functionally involved in ALS.

5. Conclusion

We performed gene expression analyses in ALS using two approaches: one that compared expression between ALS cases and controls regardless of disease severity, and one that examined changes in expression across increasing severity-grades of affected tissue. Both approaches revealed significant changes in expression for known ALS genes and highlighted several interesting candidate genes. However the approach that examined changes in expression across different grades of severity identified genes and pathways not identified by the other method. These included changes in *C9orf72*, *MATR3* and *VEGFA* expression, known ALS genes, and changes in RNA processing and immune-response pathways, known to be important in ALS. We therefore suggest that stratification by disease severity can complement ALS gene expression studies, and identify risk genes in which standard tissue-specific comparison models cannot.

6. Acknowledgements

This work was funded by a Medical Research Council PhD studentship awarded through the MRC Centre for Neurodegeneration Research, by a grant from the Motor Neuron Disease Association of Great Britain and Northern Ireland, and by the European Community's Health Seventh Framework Programme (FP7/2007–2013) 259867, and grant (075615/Z/04/z) from the Wellcome Trust. This project was supported by ZonMW, under the frame of E-Rare-2, the ERA-Net for Research on Rare Diseases. This is an EU Joint Programme - Neurodegenerative Disease Research (JPND) project. The project is supported through the following funding organisations under the aegis of JPND - www.jpnd.eu (United Kingdom, Medical Research Council, and Netherlands, ZonMW).

7. Disclosure Statement

The authors thank the ALS Association, the Angel Fund, and the ALS Therapy Alliance for support. A.A.C., C.E.S. and C.M.L. receive salary support. from the NIHR Biomedical Research Centre in Mental Health and the Biomedical Research Unit in Dementia, both at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. Tissue was provided by the London Neurodegenerative Diseases Brain Bank. The Brain Bank receives funding from the MRC and through the Brains for Dementia Research project.

8. References

- Abel, O., Powell, J.F., Andersen, P.M., Al-Chalabi, A. 2012. ALSod: A user-friendly online bioinformatics tool for amyotrophic lateral sclerosis genetics. *Human Mutation* 33(9), 1345-51.
- Al-Chalabi, A., Calvo, A., Chio, A., Colville, S., Ellis, C.M., Hardiman, O., Heverin, M., Howard, R.S., Huisman, M.H.B., Keren, N., Leigh, P.N., Mazzini, L., Mora, G., Orrell, R.W., Rooney, J., Scott, K.M., Scotton, W.J., Seelen, M., Shaw, C.E., Sidle, K.S., Swingler, R., Tsuda, M., Veldink, J.H., Visser, A.E., van den Berg, L.H., Pearce, N. 2014. Analysis of amyotrophic lateral sclerosis as a multistep process: a population-based modelling study. *The Lancet Neurology* 13(11), 1108-13.
- Al-Chalabi, A., Hardiman, O. 2013. The epidemiology of ALS: a conspiracy of genes, environment and time. *Nat Rev Neurol* 9(11), 617-28.
- Andersen, P.M., Al-Chalabi, A. 2011. Clinical genetics of amyotrophic lateral sclerosis: what do we really know? *Nat Rev Neurol* 7(11), 603-15.
- Belzil, V.V., Bauer, P.O., Prudencio, M., Gendron, T.F., Stetler, C.T., Yan, I.K., Pregent, L., Daugherty, L., Baker, M.C., Rademakers, R., Boylan, K., Patel, T.C., Dickson, D.W., Petrucelli, L. 2013. Reduced C9orf72 gene expression in c9FTD/ALS is caused by histone trimethylation, an epigenetic event detectable in blood. *Acta Neuropathol* 126(6), 895-905.
- Bernardini, C., Censi, F., Lattanzi, W., Barba, M., Calcagnini, G., Giuliani, A., Tasca, G., Sabatelli, M., Ricci, E., Michetti, F. 2013. Mitochondrial Network Genes in the Skeletal Muscle of Amyotrophic Lateral Sclerosis Patients. *PLoS ONE* 8(2), e57739.
- Bristol, L.A., Rothstein, J.D. 1996. Glutamate transporter gene expression in amyotrophic lateral sclerosis motor cortex. *Annals of Neurology* 39(5), 676-9.
- Bruneteau, G., Simonet, T., Bauché, S., Mandjee, N., Malfatti, E., Girard, E., Tanguy, M.-L., Behin, A., Khiami, F., Sariali, E., Hell-Remy, C., Salachas, F., Pradat, P.-F., Fournier, E., Lacomblez, L., Koenig, J., Romero, N.B., Fontaine, B., Meininger, V., Schaeffer, L., Hantai, D. 2013. Muscle

histone deacetylase 4 upregulation in amyotrophic lateral sclerosis: potential role in reinnervation ability and disease progression. *Brain*.

- Chiò, A., Schymick, J.C., Restagno, G., Scholz, S.W., Lombardo, F., Lai, S.-L., Mora, G., Fung, H.-C., Britton, A., Arepalli, S., Gibbs, J.R., Nalls, M., Berger, S., Kwee, L.C., Oddone, E.Z., Ding, J., Crews, C., Rafferty, I., Washecka, N., Hernandez, D., Ferrucci, L., Bandinelli, S., Guralnik, J., Macchiardi, F., Torri, F., Lupoli, S., Chanock, S.J., Thomas, G., Hunter, D.J., Gieger, C., Wichmann, H.E., Calvo, A., Mutani, R., Battistini, S., Giannini, F., Caponnetto, C., Mancardi, G.L., La Bella, V., Valentino, F., Monsurrò, M.R., Tedeschi, G., Marinou, K., Sabatelli, M., Conte, A., Mandrioli, J., Sola, P., Salvi, F., Bartolomei, I., Siciliano, G., Carlesi, C., Orrell, R.W., Talbot, K., Simmons, Z., Connor, J., Pioro, E.P., Dunkley, T., Stephan, D.A., Kasperaviciute, D., Fisher, E.M., Jabonka, S., Sendtner, M., Beck, M., Bruijn, L., Rothstein, J., Schmidt, S., Singleton, A., Hardy, J., Traynor, B.J. 2009. A two-stage genome-wide association study of sporadic amyotrophic lateral sclerosis. *Human Molecular Genetics* 18(8), 1524-32.
- Ciura, S., Lattante, S., Le Ber, I., Latouche, M., Tostivint, H., Brice, A., Kabashi, E. 2013. Loss of function of C9orf72 causes motor deficits in a zebrafish model of amyotrophic lateral sclerosis. *Ann Neurol* 74(2), 180-7.
- Daoud, H., Belzil, V., Desjarlais, A., Camu, W., Dion, P., Rouleau, G. 2010. Analysis of the UNC13A gene as a risk factor for sporadic amyotrophic lateral sclerosis. *Arch Neurol* 67(4), 516 - 7.
- Daoud, H., Valdmanis, P.N., Gros-Louis, F., et al. 2011. REsequencing of 29 candidate genes in patients with familial and sporadic amyotrophic lateral sclerosis. *Archives of Neurology* 68(5), 587-93.
- DeJesus-Hernandez, M., Mackenzie, Ian R., Boeve, Bradley F., Boxer, Adam L., Baker, M., Rutherford, Nicola J., Nicholson, Alexandra M., Finch, NiCole A., Flynn, H., Adamson, J., Kouri, N., Wojtas, A., Sengdy, P., Hsiung, G.-Yuek R., Karydas, A., Seeley, William W., Josephs, Keith A., Coppola, G., Geschwind, Daniel H., Wszolek, Zbigniew K., Feldman, H., Knopman, David S., Petersen,

Spinal Cord Gene Expression in ALS

- Ronald C., Miller, Bruce L., Dickson, Dennis W., Boylan, Kevin B., Graff-Radford, Neill R., Rademakers, R. 2011. Expanded GGGCC Hexanucleotide Repeat in Noncoding Region of C9ORF72 Causes Chromosome 9p-Linked FTD and ALS. *Neuron* 72(2), 245-56.
- Donnelly, C.J., Zhang, P.W., Pham, J.T., Haeusler, A.R., Mistry, N.A., Vidensky, S., Daley, E.L., Poth, E.M., Hoover, B., Fines, D.M., Maragakis, N., Tienari, P.J., Petrucelli, L., Traynor, B.J., Wang, J., Rigo, F., Bennett, C.F., Blackshaw, S., Sattler, R., Rothstein, J.D. 2013. RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention. *Neuron* 80(2), 415-28.
- Du, P., Kibbe, W.A., Lin, S.M. 2008. lumi: a pipeline for processing Illumina microarray. *Bioinformatics* 24(13), 1547-8.
- Futschik, M. 2012. Mfuzz: Soft clustering of time series gene expression data. R package version 2.22.0.
- Huang da, W., Sherman, B.T., Lempicki, R.A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nature protocols* 4(1), 44-57.
- Jensen, L.J., Kuhn, M., Stark, M., Chaffron, S., Creevey, C., Muller, J., Doerks, T., Julien, P., Roth, A., Simonovic, M., Bork, P., von Mering, C. 2009. STRING 8--a global view on proteins and their functional interactions in 630 organisms. *Nucleic Acids Res* 37(Database issue), D412-6.
- Johnson, J.O., Piro, E.P., Boehringer, A., Chia, R., Feit, H., Renton, A.E., Pliner, H.A., Abramzon, Y., Marangi, G., Winborn, B.J., Gibbs, J.R., Nalls, M.A., Morgan, S., Shoai, M., Hardy, J., Pittman, A., Orrell, R.W., Malaspina, A., Sidle, K.C., Fratta, P., Harms, M.B., Baloh, R.H., Pestronk, A., Weihl, C.C., Rogaeva, E., Zinman, L., Drory, V.E., Borghero, G., Mora, G., Calvo, A., Rothstein, J.D., Drepper, C., Sendtner, M., Singleton, A.B., Taylor, J.P., Cookson, M.R., Restagno, G., Sabatelli, M., Bowser, R., Chio, A., Traynor, B.J. 2014. Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis. *Nat Neurosci* 17(5), 664-6.

Spinal Cord Gene Expression in ALS

- Kanai, Y., Cl  men  on, B., Simonin, A., Leuenberger, M., Lochner, M., Weisstanner, M., Hediger, M.A. 2013. The SLC1 high-affinity glutamate and neutral amino acid transporter family. *Molecular Aspects of Medicine* 34(2-3), 108-20.
- Kim, H.J., Kim, N.C., Wang, Y.D., Scarborough, E.A., Moore, J., Diaz, Z., MacLea, K.S., Freibaum, B., Li, S., Molliex, A., Kanagaraj, A.P., Carter, R., Boylan, K.B., Wojtas, A.M., Rademakers, R., Pinkus, J.L., Greenberg, S.A., Trojanowski, J.Q., Traynor, B.J., Smith, B.N., Topp, S., Gkazi, A.S., Miller, J., Shaw, C.E., Kottlors, M., Kirschner, J., Pestronk, A., Li, Y.R., Ford, A.F., Gitler, A.D., Benatar, M., King, O.D., Kimonis, V.E., Ross, E.D., Weihl, C.C., Shorter, J., Taylor, J.P. 2013. Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS. *Nature* 495(7442), 467-73.
- Lagier-Tourenne, C., Baughn, M., Rigo, F., Sun, S., Liu, P., Li, H.R., Jiang, J., Watt, A.T., Chun, S., Katz, M., Qiu, J., Sun, Y., Ling, S.C., Zhu, Q., Polymenidou, M., Drenner, K., Artates, J.W., McAlonis-Downes, M., Markmiller, S., Hutt, K.R., Pizzo, D.P., Cady, J., Harms, M.B., Baloh, R.H., Vandenberg, S.R., Yeo, G.W., Fu, X.D., Bennett, C.F., Cleveland, D.W., Ravits, J. 2013. Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration. *Proc Natl Acad Sci U S A* 110(47), E4530-9.
- Lai, T., Jabaudon, D., Molyneaux, B.J., Azim, E., Arlotta, P., Menezes, J.R., Macklis, J.D. 2008. SOX5 controls the sequential generation of distinct corticofugal neuron subtypes. *Neuron* 57(2), 232-47.
- Lambrechts, D., Storkebaum, E., Morimoto, M., Del-Favero, J., Desmet, F., Marklund, S., Wyns, S., Thijs, V., Andersson, J., van Marion, I. 2003. VEGF is a modifier of amyotrophic lateral sclerosis in mice and humans and protects motoneurons against ischemic death. *Nat Genet* 34, 383 - 94.
- Landers, J.E., Melki, J., Meininger, V., Glass, J.D., van den Berg, L.H., van Es, M.A., Sapp, P.C., van Vught, P.W., McKenna-Yasek, D.M., Blauw, H.M., Cho, T.J., Polak, M., Shi, L., Wills, A.M., Broom, W.J.,

Spinal Cord Gene Expression in ALS

- Ticozzi, N., Silani, V., Ozoguz, A., Rodriguez-Leyva, I., Veldink, J.H., Ivinson, A.J., Saris, C.G., Hosler, B.A., Barnes-Nessa, A., Couture, N., Wokke, J.H., Kwiatkowski, T.J., Jr., Ophoff, R.A., Cronin, S., Hardiman, O., Diekstra, F.P., Leigh, P.N., Shaw, C.E., Simpson, C.L., Hansen, V.K., Powell, J.F., Corcia, P., Salachas, F., Heath, S., Galan, P., Georges, F., Horvitz, H.R., Lathrop, M., Purcell, S., Al-Chalabi, A., Brown, R.H., Jr. 2009. Reduced expression of the Kinesin-Associated Protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. *Proc Natl Acad Sci U S A* 106(22), 9004-9.
- Lin, C.-L.G., Bristol, L.A., Jin, L., Dykes-Hoberg, M., Crawford, T., Clawson, L., Rothstein, J.D. 1998. Aberrant RNA Processing in a Neurodegenerative Disease: the Cause for Absent EAAT2, a Glutamate Transporter, in Amyotrophic Lateral Sclerosis. *Neuron* 20(3), 589-602.
- Majounie, E., Renton, A.E., Mok, K., Dopper, E.G.P., Waite, A., Rollinson, S., Chiò, A., Restagno, G., Nicolaou, N., Simon-Sanchez, J., van Swieten, J.C., Abramzon, Y., Johnson, J.O., Sendtner, M., Pampillet, R., Orrell, R.W., Mead, S., Sidle, K.C., Houlden, H., Rohrer, J.D., Morrison, K.E., Pall, H., Talbot, K., Ansorge, O., Hernandez, D.G., Arepalli, S., Sabatelli, M., Mora, G., Corbo, M., Giannini, F., Calvo, A., Englund, E., Borghero, G., Floris, G.L., Remes, A.M., Laaksovirta, H., McCluskey, L., Trojanowski, J.Q., Van Deerlin, V.M., Schellenberg, G.D., Nalls, M.A., Drory, V.E., Lu, C.-S., Yeh, T.-H., Ishiura, H., Takahashi, Y., Tsuji, S., Le Ber, I., Brice, A., Drepper, C., Williams, N., Kirby, J., Shaw, P., Hardy, J., Tienari, P.J., Heutink, P., Morris, H.R., Pickering-Brown, S., Traynor, B.J. 2012. Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study. *The Lancet Neurology* 11(4), 323-30.
- Meisler, M.H., Russ, C., Montgomery, K.T., Greenway, M., Ennis, S., Hardiman, O., Figlewicz, D.A., Quenneville, N.R., Conibear, E., Brown, R.H., Jr. 2008. Evaluation of the Golgi trafficking protein VPS54 (wobbler) as a candidate for ALS. *Amyotroph Lateral Scler* 9(3), 141-8.

Spinal Cord Gene Expression in ALS

- Meyer, T., Fromm, A., Munch, C., Schwalenstocker, B., Fray, A., Ince, P., Stamm, S., Gron, G., Ludolph, A., Shaw, P. 1999. The RNA of the glutamate transporter EAAT2 is variably spliced in amyotrophic lateral sclerosis and normal individuals. *J Neurol Sci* 170, 45 - 50.
- Munch, C., Sedlmeier, R., Meyer, T., Homberg, V., Sperfeld, A., Kurt, A., Prudlo, J., Peraus, G., Hanemann, C., Stumm, G., Ludolph, A. 2004. Point mutations of the p150 subunit of dynactin (DCTN1) gene in ALS. *Neurology* 63, 724 - 6.
- Oosthuysen, B., Moons, L., Storkebaum, E., Beck, H., Nuyens, D., Brusselmans, K., Dorpe, J.V., Hellings, P., Gorselink, M., Heymans, S., Theilmeier, G., Dewerchin, M., Laudénbach, V., Vermeylen, P., Raat, H., Acker, T., Vleminckx, V., Bosch, L.V.D., Cashman, N., Fujisawa, H., Drost, M.R., Sciot, R., Bruyninckx, F., Hicklin, D.J., Ince, C., Gressens, P., Lupu, F., Plate, K.H., Robberecht, W., Herbert, J.-M., Collen, D., Carmeliet, P. 2001. Deletion of the hypoxia-response element in the vascular endothelial growth factor promoter causes motor neuron degeneration. *Nat Genet* 28(2), 131-8.
- Puls, I., Jonnakuty, C., LaMonte, B., Holzbaur, E., Tokito, M., Mann, E., Floeter, M., Bidus, K., Drayna, D., Oh, S. 2003. Mutant dynactin in motor neuron disease. *Nat Genet* 33, 455 - 6.
- Ravits, J., Appel, S., Baloh, R.H., Barohn, R., Rix Brooks, B., Elman, L., Floeter, M.K., Henderson, C., Lomen-Hoerth, C., Macklis, J.D., McCluskey, L., Mitsumoto, H., Przedborski, S., Rothstein, J., Trojanowski, J.Q., van den Berg, L.H., Ringel, S. 2013. Deciphering amyotrophic lateral sclerosis: What phenotype, neuropathology and genetics are telling us about pathogenesis. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration* 14(S1), 5-18.
- Ravits, J., Laurie, P., Fan, Y., Moore, D.H. 2007. Implications of ALS focality: Rostral–caudal distribution of lower motor neuron loss postmortem. *Neurology* 68(19), 1576-82.
- Renton, Alan E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, Jennifer C., Laaksovirta, H., van Swieten, John C., Myllykangas, L., Kalimo, H., Paetau, A.,

Spinal Cord Gene Expression in ALS

- Abramzon, Y., Remes, Anne M., Kaganovich, A., Scholz, Sonja W., Duckworth, J., Ding, J., Harmer, Daniel W., Hernandez, Dena G., Johnson, Janel O., Mok, K., Ryten, M., Trabzuni, D., Guerreiro, Rita J., Orrell, Richard W., Neal, J., Murray, A., Pearson, J., Jansen, Iris E., Sondervan, D., Seelaar, H., Blake, D., Young, K., Halliwell, N., Callister, Janis B., Toulson, G., Richardson, A., Gerhard, A., Snowden, J., Mann, D., Neary, D., Nalls, Michael A., Peuralinna, T., Jansson, L., Isoviita, V.-M., Kaivorinne, A.-L., Hölttä-Vuori, M., Ikonen, E., Sulkava, R., Benatar, M., Wu, J., Chiò, A., Restagno, G., Borghero, G., Sabatelli, M., Heckerman, D., Rogava, E., Zinman, L., Rothstein, Jeffrey D., Sendtner, M., Drepper, C., Eichler, Evan E., Alkan, C., Abdullaev, Z., Pack, Svetlana D., Dutra, A., Pak, E., Hardy, J., Singleton, A., Williams, Nigel M., Heutink, P., Pickering-Brown, S., Morris, Huw R., Tienari, Pentti J., Traynor, Bryan J. 2011. A Hexanucleotide Repeat Expansion in C9ORF72 Is the Cause of Chromosome 9p21-Linked ALS-FTD. *Neuron* 72(2), 257-68.
- Roche, J.C., Rojas-Garcia, R., Scott, K.M., Scotton, W., Ellis, C.E., Burman, R., Wijesekera, L., Turner, M.R., Leigh, P.N., Shaw, C.E., Al-Chalabi, A. 2012. A proposed staging system for amyotrophic lateral sclerosis. *Brain* 135(3), 847-52.
- Saeed, M., Siddique, N., Hung, W.Y., Usacheva, E., Liu, E., Sufit, R.L., Heller, S.L., Haines, J.L., Pericak-Vance, M., Siddique, T. 2006. Paraoxonase cluster polymorphisms are associated with sporadic ALS. *Neurology* 67(5), 771-6.
- Saunderson, R., Yu, B., Trent, R.J., Pamphlett, R. 2004. A polymorphism in the poliovirus receptor gene differs in motor neuron disease. *Neuroreport* 15(2), 383-6.
- Schymick, J.C., Scholz, S.W., Fung, H.-C., Britton, A., Arepalli, S., Gibbs, J.R., Lombardo, F., Matarin, M., Kasperaviciute, D., Hernandez, D.G., Crews, C., Bruijn, L., Rothstein, J., Mora, G., Restagno, G., Chiò, A., Singleton, A., Hardy, J., Traynor, B.J. 2007. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. *The Lancet Neurology* 6(4), 322-8.

- Shatunov, A., Mok, K., Newhouse, S., Weale, M.E., Smith, B., Vance, C., Johnson, L., Veldink, J.H., van Es, M.A., van den Berg, L.H., Robberecht, W., Van Damme, P., Hardiman, O., Farmer, A.E., Lewis, C.M., Butler, A.W., Abel, O., Andersen, P.M., Fogh, I., Silani, V., Chiò, A., Traynor, B.J., Melki, J., Meininger, V., Landers, J.E., McGuffin, P., Glass, J.D., Pall, H., Leigh, P.N., Hardy, J., Brown Jr, R.H., Powell, J.F., Orrell, R.W., Morrison, K.E., Shaw, P.J., Shaw, C.E., Al-Chalabi, A. 2010. Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *The Lancet Neurology* 9(10), 986-94.
- Shoichet, S.A., Waibel, S., Endruhn, S., Sperfeld, A.D., Vorwerk, B., Muller, I., Erdogan, F., Ludolph, A.C., Ropers, H.H., Ullmann, R. 2009. Identification of candidate genes for sporadic amyotrophic lateral sclerosis by array comparative genomic hybridization. *Amyotroph Lateral Scler* 10(3), 162-9.
- van Es, M., Veldink, J., Saris, C., Blauw, H., van Vught, P., Birve, A., Lemmens, R., Schelhaas, H., Groen, E., Huisman, M., van der Kooi, A., de Visser, M., Dahlberg, C., Estrada, K., Rivadeneira, F., Hofman, A., Zwarts, M., van Doormaal, P., Rujescu, D., Strengman, E., Giegling, I., Muglia, P., Tomik, B., Slowik, A., Uitterlinden, A., Hendrich, C., Waibel, S., Meyer, T., Ludolph, A., Glass, J. 2009. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet* 41(10), 1083 - 7.
- Warde-Farley, D., Donaldson, S.L., Comes, O., Zuberi, K., Badrawi, R., Chao, P., Franz, M., Grouios, C., Kazi, F., Lopes, C.T., Maitland, A., Mostafavi, S., Montojo, J., Shao, Q., Wright, G., Bader, G.D., Morris, Q. 2010. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Research* 38(suppl 2), W214-W20.
- Warnes, G., Bolker, B., Lumley, T. 2009. *gplots: Various R programming tools for plotting data*. R package version 2.6.0.