



Cerebrospinal Fluid Biomarker Development in Amyotrophic Lateral Sclerosis

by

Alexander Guy Thompson

A thesis submitted to the University of Oxford for the degree of
DOCTOR OF PHILOSOPHY

St John's College
Nuffield Department of Clinical Neurosciences
University of Oxford

Cerebrospinal Fluid Biomarker Development in Amyotrophic Lateral Sclerosis

Submitted for the degree of Doctor of Philosophy at the University of Oxford

Alexander G Thompson, St John's College, Michaelmas Term 2018

Abstract

Amyotrophic lateral sclerosis (ALS), the commonest form of motor neuron disease (MND), is a fatal, typically rapidly-progressive neurodegenerative syndrome without highly-effective disease-modifying treatment. A diverse range of upstream cellular perturbations converge on a common clinical pathway. A major impediment to the development of efficacious treatment is the lack of sensitive, objective measures of the disease process to enable early diagnosis, stratification and monitoring necessary for therapeutic trials. The experiments described in this thesis aim to develop novel biomarkers using mass spectrometry-based proteomic analysis of cerebrospinal fluid (CSF) and, to provide improved access to the intracellular compartment of central nervous system (CNS) cells, CSF-derived extracellular vesicles (EVs).

A high-throughput proteomic analysis of individual whole CSF samples from patients with ALS and healthy controls, patients with ALS-mimicking conditions and the upper motor neuron only variant primary lateral sclerosis (PLS) was performed. This analysis demonstrated elevation of three chitinase proteins, Chitotriosidase-1 (CHIT1), Chitinase-3-like protein 1 (CHI3L1) and Chitinase-3-like protein 2 (CHI3L2) in the CSF of ALS patients. Chitinase proteins correlated with the rate of progression of the disease and the marker of axonal degeneration phosphorylated neurofilament heavy chain (pNFH) and changed little over time in longitudinal analysis.

The elevation of chitinase proteins was validated using a commercially-available enzyme-linked immunosorbent assay (ELISA). Additional analysis of chitinase proteins and clinical parameters identified correlations between the burden of upper motor neuron (UMN) signs and degree of cognitive dysfunction with CHI3L1, whilst CHI3L2 and CHIT1 correlated with the rate of symptom progression and axonal degeneration as measured by pNFH. This provides evidence that the different chitinase proteins reflect different dimensions of the glial response in ALS and supports their use as a measure of target engagement in trials of therapies directed at non-cell-autonomous mechanisms.

In order to enhance access to the intracellular proteome, a method of extraction of EVs from CSF based on size-exclusion chromatography was employed. Comparison of this method with current standard methods showed improvement in EV recovery and proteomic depth. Enrichment analysis comparing whole CSF and CSF EV-enriched proteomes suggested that the EV proteome reflects intracellular processes, contrasting with the secreted and plasma-derived proteome of whole CSF. Analysis of open access datasets suggests a predominantly white matter and choroid plexus origin of the EV proteome and additional overlap with the proteome of cortical and hippocampal neurons.

This optimised EV extraction method was applied to individual CSF samples from a cohort of ALS patients and controls. This identified alterations in the levels of proteins involved in proteostasis pathways in ALS CSF, in particular the hexameric proteasome-like protease bleomycin hydrolase.

In conclusion, the work described in this thesis suggest that the whole CSF proteome in ALS predominantly reflects a secretory glial response and provides promising biomarkers of glial activity in ALS, whilst the CSF EV proteome is more reflective of core intracellular events implicated in ALS.

Acknowledgements

I am indebted to my supervisors. I would like to thank Professor Martin Turner, without whom I would not have begun this interesting and challenging project. I am extremely grateful to Professor Turner for providing the opportunity and inspiration to study ALS, for imparting his breadth of understanding of this multifaceted disease, his unwavering support and guidance through the melee of academia. To Professor Kevin Talbot I am grateful for his enthusiasm and encouragement in supporting the development of this project, his excellent careers advice and belief in my ability.

I would like to thank Dr Elizabeth Gray for her enormous contribution to all aspects of this project. As well as teaching me virtually all of my laboratory skills, Liz always remained upbeat and provided encouragement when things were hard, as well as a dose of realism, tempering my attempts to try every permutation of a protocol (and my inexplicable belief in my mental arithmetic skills). The success of both the whole CSF and CSF EV work and my development as a scientist owe a great deal to Liz's efforts.

I would also like to thank several members of the Target Discovery Institute, specifically Marie-Läetitia Thézénas and Roman Fischer for their help with the optimisation and execution of whole CSF and EV proteomic analysis, Philip Charles for his advice on the analysis of the resulting proteomic datasets, and Professor Benedikt Kessler for his support and advice throughout the project.

I am grateful to Professor Matthew Wood and the members of his group for sharing their EV extraction methodology, lab space and instrumentation. Particular thanks go to Dr Imre Mager for his instruction in the art of size exclusion chromatography, nanoparticle tracking analysis and help establishing the method for extraction of CSF EVs.

Thank you also to Alexander Bampton, for performing the CSF chitotriosidase activity assays, to Dominika Raciborska for her contribution to the C-reactive protein (CRP) ELISA and again to Liz for the CSF pNFH measurements. I am also grateful to all other members of the Oxford MND Clinic and research team for making the last 3 $\frac{1}{2}$ years into an experience that I will cherish: in particular Malcolm Proudfoot for his early guidance, Jakub Scaber, Emily Feneberg and Lucy Farrimond for friendship, lively and entertaining debate and the odd lumbar puncture, and Lynn Ossher, who overcame my woeful lack of organisational skills.

I am grateful to my parents for their unfaltering encouragement, patience and support (financial and otherwise) over the last 35 years, and to my wonderful wife, Priya, for her espousal of my work as well as myself.

Finally, I would like to thank the patients and volunteers who contributed a great deal of time and effort, often undergoing the discomfort of multiple lumbar punctures for no personal benefit. They are the embodiment of dignity in the face of great adversity.

For *Priya*, *Kiran* and *Rohan*.

Contents

Abstract	i
Acknowledgements	i
Contents	iii
List of figures	x
List of tables	xiii
Abbreviations	xvi
1 Introduction	1
1.1 Amyotrophic lateral sclerosis	1
1.2 Amyotrophic Lateral Sclerosis	2
1.2.1 Epidemiology of ALS	2
1.2.2 Clinical features of ALS	3
1.2.3 Natural history and prognosis of ALS	7
1.2.4 Neuropathology of ALS	8
1.2.5 Genetics of ALS	9
1.2.6 Pathomechanisms in ALS	14

1.3	Biomarkers of ALS	25
1.3.1	Cerebrospinal fluid biomarkers of ALS	30
1.3.2	Extracellular vesicles	41
1.4	Aims of the work presented in this thesis	44
2	Materials and Methods	45
2.1	Proteomic analysis	45
2.1.1	Proteomic analysis in biomarker development	46
2.1.2	LC-MS/MS	47
2.1.3	Data analysis	57
2.2	Size exclusion chromatography	60
2.2.1	Basic principles of size exclusion chromatography	60
2.3	Participants and sampling	61
2.3.1	Ethical Approval and Subject Recruitment	61
2.3.2	Biofluid Sampling and Processing	62
2.3.3	Clinical parameters	63
2.4	Statistical analysis	64
3	The Cerebrospinal Fluid Proteome in ALS	65
3.1	Hypothesis and Aims of This Chapter	65
3.1.1	Hypothesis	65

3.1.2	Aims	65
3.2	Introduction	66
3.3	Materials and methods	67
3.3.1	Participants and sampling	67
3.3.2	Sample preparation for proteomic analysis	67
3.3.3	Proteomic analysis	68
3.3.4	ELISA	69
3.3.5	Statistical analysis	70
3.4	Results	71
3.4.1	Proteomic analysis performance	71
3.4.2	Cross-sectional candidate biomarker identification	72
3.4.3	Classifier performance of identified candidate markers	74
3.4.4	Cross-sectional linear regression models of chitinase proteins and clinical parameters	78
3.4.5	Longitudinal analysis of chitinase proteins	81
3.4.6	Survival analysis	81
3.4.7	Validation of previously identified candidate markers	82
3.4.8	Discussion	85
4	Validation of Chitinase Biomarkers in ALS	88

4.1	Hypothesis and Aims of This Chapter	88
4.1.1	Hypothesis	88
4.1.2	Aims	88
4.2	Introduction	89
4.3	Materials and methods	90
4.3.1	Participants and sampling	90
4.3.2	ELISA and CHIT1 activity assay	92
4.3.3	CHIT1 activity measurement	95
4.3.4	Statistical analysis	95
4.4	Results	96
4.4.1	Demographic data	96
4.4.2	Correlation of ELISA and proteomic data	96
4.4.3	Cross-sectional chitinase protein levels	97
4.4.4	Classifier performance	99
4.4.5	Correlation with clinical parameters	100
4.4.6	Correlation of biochemical parameters	103
4.4.7	Longitudinal analysis	106
4.4.8	Survival analysis	106
4.4.9	Serum chitinase proteins	108

4.5	Discussion	109
5	UFLC-derived CSF Extracellular Vesicle Origin and Proteome	113
5.1	Hypothesis and Aims of This Chapter	113
5.1.1	Hypothesis	113
5.1.2	Aims	113
5.2	Introduction	113
5.3	Materials and Methods	115
5.3.1	Participants and Sampling	115
5.3.2	EV extraction	115
5.3.3	Ultracentrifugation	116
5.3.4	Ultrafiltration-liquid chromatography	116
5.3.5	EV characterisation	117
5.3.6	Immunoblotting	117
5.3.7	Transmission electron microscopy	118
5.3.8	Sample preparation for proteomic analysis	118
5.3.9	Liquid chromatography-tandem mass spectrometry	119
5.3.10	Normalisation and statistical analysis	120
5.4	Results	122
5.4.1	Characterisation of UFLC and UC performance	122

5.4.2	The CSF-EV proteome	127
5.5	Discussion	133
6	The Cerebrospinal Fluid Extracellular Vesicle Proteome in ALS	137
6.1	Hypothesis and aims	137
6.1.1	Hypothesis	137
6.1.2	Aims	137
6.2	Introduction	138
6.3	Materials and methods	139
6.3.1	Participants and sampling	139
6.3.2	EV extraction	140
6.3.3	EV characterisation	140
6.3.4	Sample preparation for proteomic analysis	140
6.3.5	Liquid chromatography-tandem mass spectrometry	141
6.3.6	Normalisation and statistical analysis	142
6.4	Results	143
6.4.1	EV size distribution and number	143
6.4.2	The CSF EV proteome in ALS	143
6.5	Discussion	151
7	Discussion	156

7.1	The ALS CSF proteome and the chitinase proteins	156
7.2	The CSF EV proteome	159
7.3	Future directions	160
7.4	Closing remarks	165
8	Publications arising from this research period	166
8.1	Original articles	166
8.2	Reviews	167
	References	168

List of Figures

1.1	Age of onset of ALS	3
1.2	Survival of patients with ALS	8
1.3	Frequency of major ALS genes	10
1.4	Proposed cellular mechanisms of ALS	15
1.5	EV biogenesis	42
2.1	Linear Quadrupole	51
2.2	Schematic of the Q Exactive TM HF Hybrid Quadrupole-Orbitrap TM mass spectrometer	51
2.3	Fragmentation of precursor ions	53
2.4	Inference of amino acid sequence by MS2 spectrum	54
2.5	Parallel reaction monitoring	56
3.1	Cross-sectional proteomic analysis	73
3.2	Cross-sectional chitinase levels in ALS and controls	74
3.3	Classifier performance of chitinase proteins	76
3.4	Linear regression of chitinase proteins with clinical parameters	80

3.5	Longitudinal stability of Chitinase Proteins	82
3.6	Survival Analysis of Chitinase Proteins	83
3.7	Additional Candidate Validation	84
4.1	Correlation of proteomic and ELISA measurement of chitinase proteins	96
4.2	Cross-sectional CSF chitinase levels in ALS and control subjects . . .	98
4.3	Classifier performance of chitinase proteins	99
4.4	Chitinase protein correlation with clinical parameters	101
4.5	Correlation between CSF chitinase proteins and chitotriosidase activity	104
4.6	Correlation between CSF chitinase proteins and pNFH levels	105
4.7	Correlation between CSF chitinase proteins and CRP levels	105
4.8	Longitudinal analysis of chitinase proteins	106
4.9	Survival associations of chitinase proteins	107
4.10	Serum chitinase proteins in ALS and healthy controls	108
5.1	Extraction of EVs from CSF by UFLC	123
5.2	Comparison of UC and UFLC	125
5.3	EV and CSF proteomes compared with vesicle protein databases . . .	128
5.4	GO enrichment analysis of whole CSF and CSF EV-enriched proteomes	129
5.5	Origin of CSF EVs	131

6.1	Size distribution of CSF EVs in ALS and controls	144
6.2	Size distribution of CSF EVs in <i>C9</i> and non- <i>C9</i> ALS	145
6.3	Heatmap and PCA biplot of CSF EV proteome in ALS	147
6.4	Volcano plot of the CSF EV proteome in ALS	148
6.5	GO enrichment analysis of the CSF EV proteome in ALS	154
6.6	Comparison of <i>C9orf72</i> -associated and non- <i>C9</i> ALS CSF EV proteomes	155

List of Tables

1.1	Proteomic biomarker candidates in ALS	35
3.1	Whole CSF analysis demographic data	68
3.2	Relative abundance of chitinase proteins	72
3.3	Classifier performance of chitinase proteins	75
3.4	Multiple linear regression models of disease progresison rate	81
3.5	Cox proportional hazards survival modelling	83
4.1	Demographic data for the CSF ELISA cohort	91
4.2	Classifier performance of chitinase proteins	100
4.3	Linear models of disease progression rate	102
4.4	Linear models of UMN score	102
4.5	Linear models of cognitive dysfunction	103
4.6	Cox proportional hazards modelling	107
6.1	Demographic data for the CSF EV study	139
6.2	Demographic data for CSF EV proteomic analysis	146

6.3	Relative abundance of proteins FDR-adjusted $p < 0.1$	149
-----	---	-----

Abbreviations

ACTN4 Alpha-actinin-4.

ALS amyotrophic lateral sclerosis.

ALSFRS-R revised ALS functional rating scale.

ANXA11 Annexin 11.

ARF6 ADP-ribosylation factor 6.

ASO antisense oligonucleotide.

AUC area under the curve.

BH Benjamini-Hochberg step-up procedure.

BLMH Bleomycin hydrolase.

CCL3 C-C motif chemokine 3, also known as monocyte inflammatory protein 1 α .

CFL1 Cofilin-1.

CHFR Checkpoint with forkhead and ring finger domains.

CHI3L1 Chitinase-3-like protein 1.

CHI3L2 Chitinase-3-like protein 2.

CHIT1 Chitotriosidase-1.

CI confidence interval.

CID collision-induced dissociation.

CJD Creutzfeldt-Jakob disease.

CK Creatinine kinase.

CNS central nervous system.

CRP C-reactive protein.

CSF cerebrospinal fluid.

CSF1R Macrophage colony-stimulating factor 1 receptor.

CSF2 Granulocyte-macrophage colony-stimulating factor.

CSF3 Granulocyte colony-stimulating factor.

CV coefficient of variation.

CYBB Cytochrome b-245 heavy chain.

DDA data-dependent acquisition.

DIA data-independent acquisition.

DMSO dimethyl sulfoxide.

DNA deoxyribonucleic acid.

DTT dithiothreitol.

EAAT2 Excitatory amino acid transporter 2.

ECAS Edinburgh cognitive and behavioural ALS screen.

ELISA enzyme-linked immunosorbent assay.

EM electron microscopy.

ER endoplasmic reticulum.

ESCRT endosomal sorting complex required for transport.

ESI electrospray ionisation.

EV extracellular vesicle.

FDR false discovery rate.

FGF2 Fibroblast growth factor 2, also known as FGF basic.

FLAIR fluid attenuated inversion recovery.

FTD frontotemporal dementia.

FTICR Fourier transform ion cyclotron resonance.

FTL Ferritin light chain.

FUS RNA-binding protein FUS.

FVC forced vital capacity.

FWER family-wise error rate.

FWHM full-width at half maximum.

FYCO1 FYVE and coiled-coil domain-containing protein 1.

GO gene ontology.

GPMB Transmembrane glycoprotein NMB.

HCD higher-energy collisional dissociation.

HDL high-density lipoprotein.

HGNC HUGO Gene Nomenclature Committee.

HHD hand-held dynamometry.

HPLC high-performance liquid chromatography.

HR hazard ratio.

HRP horseradish peroxidase.

HUWE1 HECT, UBA and WWE domain containing 1 E3 ubiquitin protein ligase.

IAA iodoacetamide.

IFNG Interferon gamma.

IL Interleukin.

ILV intraluminal vesicle.

iPSC induced pluripotent stem cell.

IQR interquartile range.

k NN k -nearest neighbours.

L1CAM L-1 cell adhesion molecule.

LAMP2 Lysosome-associated membrane glycoprotein 2.

LC liquid chromatography.

LCM laser capture microdissection.

LC-MS/MS liquid chromatography-tandem mass spectrometry.

LDL low-density lipoprotein.

LDS lithium dodecyl sulphate.

LLOQ lower limit of quantitation.

LMN lower motor neuron.

m/z mass-to-charge ratio.

MAD median absolute deviation.

MALDI-TOF matrix-assisted laser desorption-ionisation-time of flight.

MAR missing at random.

MCP1 Monocyte chemoattractant protein-1.

MCT1 monocarboxylic acid transporter 1.

miRNA micro-RNA.

MLE maximum likelihood estimator.

MMN multifocal motor neuropathy with conduction block.

MMT manual muscle testing.

MNAR missing not at random.

MND motor neuron disease.

MRI magnetic resonance imaging.

MRM multiple reaction monitoring.

mRNA messenger RNA.

MS mass spectrometry.

MS/MS tandem mass spectrometry.

MSP multisystem proteinopathy.

MV microvesicle.

MVB multivesicular body.

MVIC maximum voluntary isometric contraction.

MWCO molecular weight cut-off.

NFL neurofilament light chain.

NLS nuclear localisation sequence.

NMR nuclear magnetic resonance.

NPEPPS Puromycin-sensitive aminopeptidase.

NTA nanoparticle tracking analysis.

nUHPLC LC-MS/MS nano ultra-high performance liquid chromatography tandem mass spectrometry.

OR odds ratio.

p75NTR^{ECD} p75 neurotrophin receptor extracellular domain.

PBS phosphate buffered saline.

PCA principal component analysis.

PDCD6IP Programmed cell death 6-interacting protein, also known as ALG-interacting protein X (ALIX).

PET positron emission tomography.

PLS primary lateral sclerosis.

PLS-DA partial least squares discriminant analysis.

PMA progressive muscular atrophy.

pNFH phosphorylated neurofilament heavy chain.

pNFH phosphorylated neurofilament heavy chain.

PPI protein-protein interaction.

PR disease progression rate.

PRM parallel reaction monitoring.

PRP^C prion protein.

PRP^{Sc} misfolded prion protein.

PSM peptide spectral match.

PTM post-translational modification.

RAB7A Ras-related protein Rab-7a.

RALB Ras-related protein Ral-B.

RAN translation repeat-associated non-ATG translation.

RF radio frequency.

RNA ribonucleic acid.

RNF213 E3 ubiquitin-protein ligase RNF213.

ROC receiver operator characteristic.

ROTS reproducibility-optimised test statistic.

RPLC reverse-phase liquid chromatography.

RRM RNA recognition motif.

SCA2 spinocerebellar ataxia type 2.

SD standard deviation.

SDCBP Syntenin-1.

SDS sodium dodecyl sulphate.

SDS-PAGE sodium dodecyl sulphate polyacrylamide gel electrophoresis.

SELDI-TOF surface-enhanced laser desorption-ionisation-time of flight.

siRNA small interfering RNA.

SOD-1 Superoxide dismutase [Cu-Zn].

SPE solid-phase extraction.

SRM selected reaction monitoring.

SRPX Sushi repeat-containing protein SRPX2.

SVD singular value decomposition.

SWATH-MS sequential window acquisition of all theoretical fragment ion spectra MS.

TDP-43 TAR DNA-binding protein 43.

TEAB triethylammonium bicarbonate.

TEM transmission electron microscopy.

TFA trifluoroacetic acid.

TGFB Transforming growth factor beta.

TGM2 Protein-glutamine gamma-glutamyltransferase 2.

Th T-helper.

Th Thomson.

TNF α Tumor necrosis factor alpha.

TOF time-of-flight.

TSG101 Tumour susceptibility gene 101 protein.

UBA1 Ubiquitin-like modifier-activating enzyme 1.

UBB Ubiquitin B.

UC ultracentrifugation.

UFLC ultrafiltration-liquid chromatography.

UHPLC ultra high performance liquid chromatography.

UMN upper motor neuron.

UPLC ultra high-performance liquid chromatography.

UPS ubiquitin-proteasome system.

VEGF Vascular endothelial growth factor.

VMP1 Vacuole membrane protein 1.

1. Introduction

1.1 Amyotrophic lateral sclerosis

Amyotrophic lateral sclerosis (ALS), the commonest form of MND, is a fatal adult-onset neurodegenerative disease characterised by unmitigating progressive muscle weakness due to degeneration of UMNs of the motor cortex and lower motor neurons (LMNs) of the brainstem and spinal cord ventral horn (Brown et al. 2017; Lowe 1994). Most cases occur sporadically, though around 5% are associated with a family history (Byrne et al. 2011; Zou et al. 2017). ALS has clinical, pathological and genetic overlap with frontotemporal dementia (FTD) (Strong et al. 2011; Ringholz et al. 2005; Phukan et al. 2012; Gijssels et al. 2012; Murray et al. 2011; Renton et al. 2011; Neumann et al. 2006). Despite two licensed disease-modifying treatments, ALS remains one of the most aggressive neurodegenerative diseases, leading to death typically within three years of symptom onset (Kiernan et al. 2011; Steenland et al. 2010).

Over 95% of cases are associated with inclusions of hyperphosphorylated, ubiquitinated, and N-terminally truncated TAR DNA-binding protein 43 (TDP-43) (Neumann et al. 2006). Despite multiple proposed mechanisms, the events leading to motor neuron death and TDP-43 deposition in ALS remain enigmatic. Current evidence points towards important roles for perturbations in RNA biology, protein degradation and quality control, cellular stress and non-cell-autonomous mechanisms (Taylor et al. 2016).

The development of effective treatment for ALS is hampered by the heterogeneity of clinical manifestations, course and survival; effective biomarkers to enable stratification, monitoring and diagnosis of patients are imperative to overcome this major

challenge (Turner et al. 2015).

This introduction will give an overview of the epidemiology, clinical and pathological features, diagnosis and genetics of ALS, and its proposed pathomechanisms. The current state-of-the-art in ALS biomarkers will also be summarised, with particular focus on CSF proteomic analysis.

1.2 Amyotrophic Lateral Sclerosis

1.2.1 Epidemiology of ALS

ALS is an uncommon disease. Its incidence has been best characterised in Europe, where estimates mostly vary between 1.47-2.43 per 100,000 persons per year (Chio et al. 2013a). Ascertainment elsewhere has been inconsistent, though estimates have been similar. Some studies have reported differences in ALS incidence according to ethnic origin, with lower incidence in people of Afrocarribean and Asian origin compared with those of European origin in the United States of America (Cronin et al. 2007).

The mean age of onset of ALS is in the seventh decade, though symptom onset before age 20 has been observed (Chio et al. 2013a; Logroscino et al. 2010, **Figure 1.1**). Incidence appears to taper beyond age 75, though whether this represents inaccurate reporting or is a function of the disease process is unclear (Logroscino et al. 2010). ALS is more common in men, with a male-to-female ratio of approximately 1.4:1, though this approaches unity at older ages (Logroscino et al. 2010; Logroscino et al. 2005; Traynor et al. 1999).

The search for other risk factors of ALS has yielded inconsistent results. Preceding head trauma, higher levels of physical activity, smoking, exposure to heavy metals, military service and exposure to pesticides have all at times been associated with

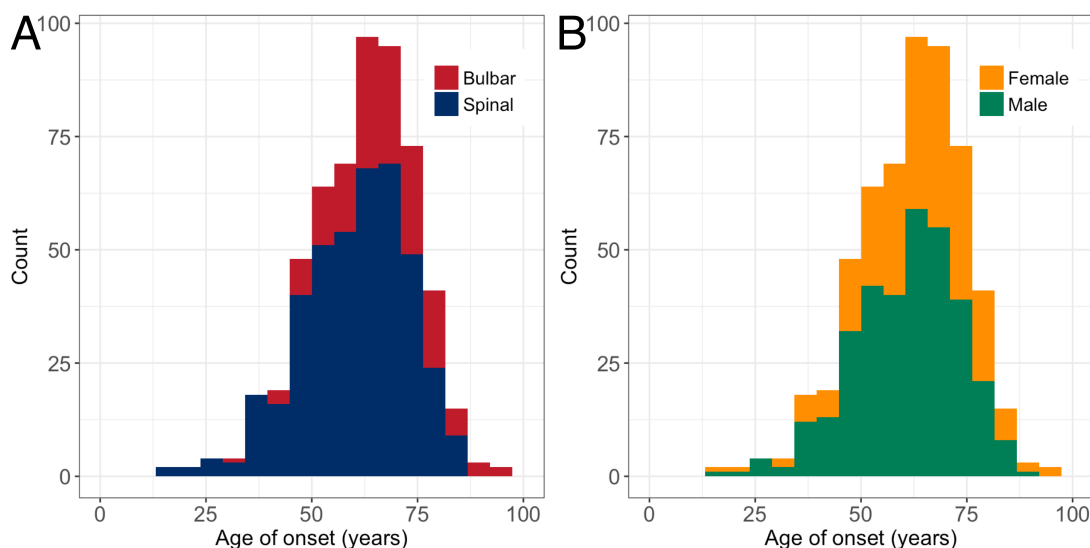


Figure 1.1: Age of symptom onset in ALS based on data from the Oxford MND Centre clinic database. Histograms illustrating age of onset counts **A** stratified by site of onset and **B** stratified by gender.

increased risk of developing ALS (Wang et al. 2017); the most consistent association appears to be that of physical activity with risk of ALS (Visser et al. 2018).

1.2.2 Clinical features of ALS

Progressive weakness is the principal presenting symptom of ALS, beginning focally in one limb in the majority of cases (over 60%), with the remainder beginning in the bulbar territory with alteration in speech and swallowing (around 30%), or rarely in the truncal musculature (limb or truncal onset together referred to as *spinal*; Chio et al. 2011; Gubbay et al. 1985). Onset site varies with age, bulbar onset being more common in older age (Chio et al. 2011, **Figure 1.1**). Muscle cramps are a common accompanying symptom (Gubbay et al. 1985). Although widespread muscle fasciculation is a frequent examination finding at presentation and may predate weakness, it is seldom the presenting complaint (Gubbay et al. 1985).

Weakness is accompanied by symptomatic frontal dysexecutive and behavioural dysfunction in under 5% of cases, although neuropsychometric testing identifies a degree

of cognitive dysfunction in up to 50% of patients (Strong et al. 2011; Ringholz et al. 2005; Phukan et al. 2012).

Mood-incongruent or excessive crying or, less frequently, laughter (known as pseudobulbar affect) is common in patients with ALS, particularly those with bulbar involvement (Gallagher 1989; Thakore et al. 2017). This occurs in effect as hyperreflexia of emotional responses due to degeneration of corticobulbar fibres (Swash et al. 2017).

Sensory symptoms, in particular pain correlating with the degree of impairment, are present in a minority of patients with ALS; paraesthesiae are rarely reported at diagnosis and are not the predominant feature (Chio et al. 2012; Gubbay et al. 1985).

Clinical examination findings in ALS reflect the involvement of the two major components of the monosynaptic corticomotoneuronal system. UMNs comprise cortical pyramidal cells principally arising in cortical layer V of the primary motor, supplementary motor and premotor cortical areas. They project axons through the corticobulbar and corticospinal tracts to the brainstem nuclei and spinal cord ventral horn (respectively), where they synapse directly with LMNs in the spinal cord ventral horn, of which α motor neurons are the main effectors of voluntary muscle contraction (reviewed in Martin 2005).

UMN dysfunction is appreciable on clinical examination as spasticity, hyperreflexia of the deep tendon reflexes and upgoing plantar responses (reflecting loss of descending inhibitory output from UMNs; Swash 2012). In ALS these may also be accompanied by other pathological reflexes such as the glabellar response and jaw jerk (Turner et al. 2013a; Strong et al. 2003). LMN dysfunction manifests as muscle wasting, fasciculation and hyporeflexia (Swash 2012; Turner et al. 2013a). The relative burden of UMN and LMN involvement on clinical examination in ALS varies widely between individuals, though establishing the extent of this is difficult since

extensive LMN dysfunction may obscure UMN signs (Swash 2012).

ALS, PLS and PMA

ALS, with clinical involvement of both UMNs and LMNs, may be separated from other forms of MND with exclusively UMN signs, known as PLS, and exclusively LMN signs, known as progressive muscular atrophy (PMA). Whether the subclassification of MND into ALS, PLS and PMA is worthwhile is a matter of debate.

PLS accounts for <5% of MND cases overall (Chio et al. 2011; Le Forestier et al. 2001). Physical signs reflect exclusive involvement of upper motor neurons, with spasticity and hyperreflexia in the absence of wasting or fasciculations (Pringle et al. 1992). The diagnosis of PLS requires the absence of physical and electrophysiological evidence of LMN dysfunction for at least four years from symptom onset (Gordon et al. 2006). Despite this delay, late onset of lower motor neuron signs (heralding ALS) is described in over 20% of patients diagnosed with PLS after four years (Gordon et al. 2006).

The relationship between PLS and ALS is debated. Patients with PLS rarely report a family history and cases of PLS within ALS pedigrees are scarce (Brugman et al. 2005). Genetic studies have suggested that ALS-associated gene mutations are rare in patients with PLS, and that mutations associated with other neurodegenerative diseases are also observed in patients with PLS (Rheenen et al. 2012; Mitsumoto et al. 2015). This suggests that, whilst a subset of PLS might represent extreme UMN-dominant ALS, the underlying cellular processes leading to neurodegeneration may differ from ALS in other cases.

The relationship between ALS and PMA is clearer. Fewer than 5% of patients present with purely LMN signs (Chio et al. 2011). Of those manifesting only LMN signs for the entire duration of the disease, around 50% have evidence of UMN involvement *post mortem* (therefore actually representing cases of ALS; Ince et al.

2003). In keeping with these findings, more extensive CNS involvement has been demonstrated *in vivo* with the use of magnetic resonance imaging (MRI), showing altered white matter integrity in PMA (Graaff et al. 2011).

Diagnosing ALS

A typical history of progressive, usually asymmetrical, weakness without significant sensory loss, accompanied by findings indicative of LMN and UMN dysfunction in the same anatomical territory, particularly when accompanied by florid fasciculations, is sufficient to make a diagnosis of ALS (Turner et al. 2013a).

Neurophysiological testing may be used as an adjunct to clinical examination to seek additional evidence of LMN dysfunction, for example identifying LMN dysfunction where only clinical UMN dysfunction is present or LMN dysfunction in a region clinically unaffected; or to identify features consistent with conditions known to present in a similar way to ALS (so-called "mimic" conditions such as multifocal motor neuropathy with conduction block (MMN) Brooks et al. 2000a; Carvalho et al. 2008; Turner et al. 2013a).

Consensus diagnostic criteria based on clinical and neurophysiological findings were initially established in 1990 and subsequently revised to improve upon the low sensitivity of early versions (Brooks 1994; Brooks et al. 2000a; Carvalho et al. 2008), particularly early in the disease and in patients with bulbar onset of symptoms (Carvalho et al. 2009; Schrooten et al. 2011). Whilst these criteria may be useful in the context of clinical research, they, along with distinctions into *clinically definite*, *clinically probable* and *clinically possible*, are of limited use to patients suffering ALS, for whom they may delay diagnosis and create uncertainty when compared to the performance of an experienced clinician (Schrooten et al. 2011).

Additional ancillary investigations are typically employed in the exclusion of mimic conditions, such as serum Creatinine kinase (CK) and muscle biopsy, where myopa-

thy is suspected; MRI in order to exclude mimics such as cervical myeloradiculopathy in upper limb-onset cases (noting that corticospinal tract hyperintensity at multiple intracranial levels using fluid attenuated inversion recovery (FLAIR) MRI sequences is observed in a proportion of patients with ALS, which might provide evidence of UMN involvement in cases with clinical evidence of LMN dysfunction clinically; Carvalho et al. 2008; Turner et al. 2013a; Hecht et al. 2001; Jin et al. 2016).

1.2.3 Natural history and prognosis of ALS

The weakness of ALS progresses both in degree at the site of onset and spatially to contiguous anatomical segments (Ravits et al. 2007; Ravits et al. 2009). The rate of accrual of disability and muscle weakness appears to change little once symptoms are established, though whether the decline is linear is controversial (Proudfoot et al. 2016; Gordon et al. 2010; Munsat et al. 1988). Muscle weakness eventually involves the diaphragm leading to progressive neuromuscular respiratory dysfunction, which is the most frequent cause of death from ALS (Corcia et al. 2008). Although in most cases death occurs within three years of symptom onset, there is wide variation in the survival of patients with ALS, ranging from months to over 10 years in up to 10% of patients (Logroscino et al. 2010; Norris et al. 1993).

Several factors have been identified that are associated with shortened or prolonged survival in ALS. Older age at symptom onset, accompanying FTD, symptom onset in the bulbar or respiratory territory, the presence of combined UMN and LMN signs, a high intrinsic rate of progression and a short latency from symptom onset to diagnosis are associated with shortened survival. Regionally-restricted forms, such as the so-called "flail-arm" variant, which may remain confined to the upper limbs for several years, and exclusively LMN or UMN signs (i.e. PLS or PMA) are associated with better prognosis (Chio et al. 2011; Logroscino et al. 2010; Westeneng et al. 2018; Knibb et al. 2016, **Figure 1.2**).

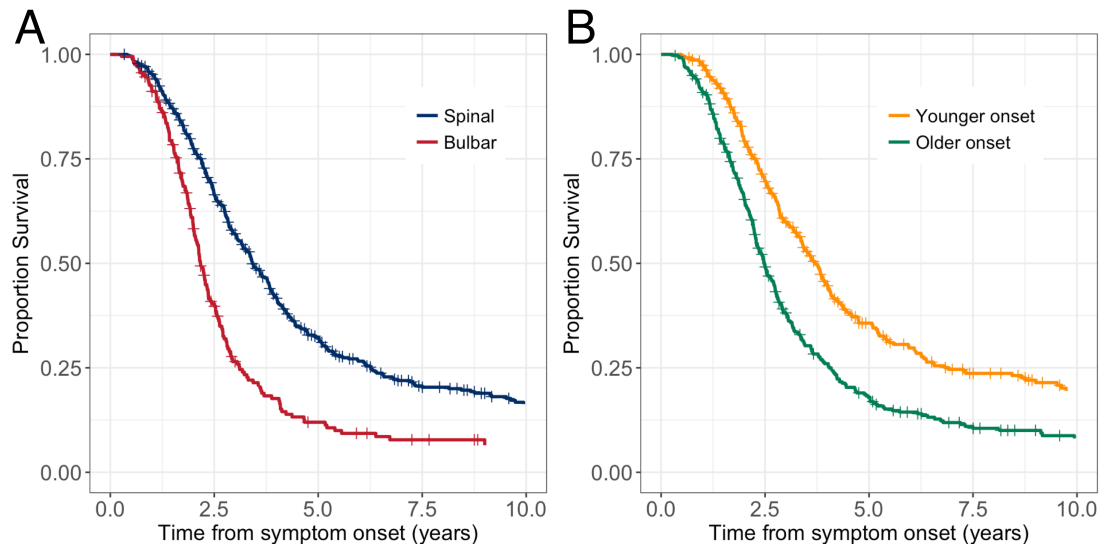


Figure 1.2: Survival of ALS patients, based on data from the Oxford MND Centre clinic database. **A** stratified by site of onset **B** stratified by age of onset (older or younger than median age of onset).

Some genetic forms of ALS are also associated with altered prognosis. Expansions in *C9orf72* are associated with shortened survival (Irwin et al. 2013). Mutations in *FUS* are also associated with shortened survival, with particularly aggressive juvenile-onset cases described (Groen et al. 2010; Baumer et al. 2010). Some *SOD1* mutations are associated with shortened survival (such as G93A and A4V mutations), whilst others are associated with longer survival (G93C and D90A mutations; Yamashita et al. 2015).

1.2.4 Neuropathology of ALS

The major pathological changes of ALS reflect the involvement of both lower and upper motor neuron populations. At the LMN level, loss of spinal cord ventral horn and brainstem motor neurons is observed, accompanied by loss of large myelinated fibres and axonal degeneration in the ventral root (Sabeti et al. 2015). Loss of layer V pyramidal cells including Betz cells of the motor cortical regions (particularly the primary motor cortex, but also the premotor cortex and supplementary motor cortex) is seen, along with degeneration of the corticospinal tracts (Sabeti et al.

2015).

Glial pathology occurs in ALS. Widespread astrogliosis and microglial reactivity has been observed in ALS, both *in vivo* and *post mortem* (Schiffer et al. 1996; Nagy et al. 1994; Kushner et al. 1991; Turner et al. 2004; Brettschneider et al. 2012).

Pathological aggregates of phosphorylated, ubiquitinated and N-terminally truncated forms of the ribonucleic acid (RNA)-binding protein TDP-43 are found in 95% of cases in cortical and ventral horn neurons in ALS (Neumann et al. 2006). This is accompanied by loss of the normal nuclear staining of TDP-43 in affected neurons (Neumann et al. 2006). TDP-43-positive inclusions are also observed in oligodendrocytes (Brettschneider et al. 2013). Similar changes are observed in affected regions in around 50% of cases of FTD (Neumann et al. 2006). Although a characteristic and widespread feature of ALS, TDP-43 inclusions may be observed in other neurological conditions including Alzheimer’s disease and chronic traumatic encephalopathy (McKee et al. 2015; Josephs et al. 2017).

The extent of *post mortem* phospho-TDP-43 pathology has been used to construct a 4-part pathological staging system for ALS pathology (Brettschneider et al. 2013; Braak et al. 2013), due to the almost heirarchical observation of the location of TDP-43 pathology in the samples studied: all cases demonstrated pathology in spinal cord and brainstem motor neurons, and agranular motor cortex; a subset of cases also demonstrated involvement of reticular formation and precerebellar nuclei; a subset of these had involvement of prefrontal cortex and basal ganglia; a subset of these also showed involvement of temporal lobe structures (Brettschneider et al. 2013).

1.2.5 Genetics of ALS

Although ALS is usually sporadic, around 5% of patients will report a family history of ALS or FTD in a first degree relative, indicating an autosomal dominantly

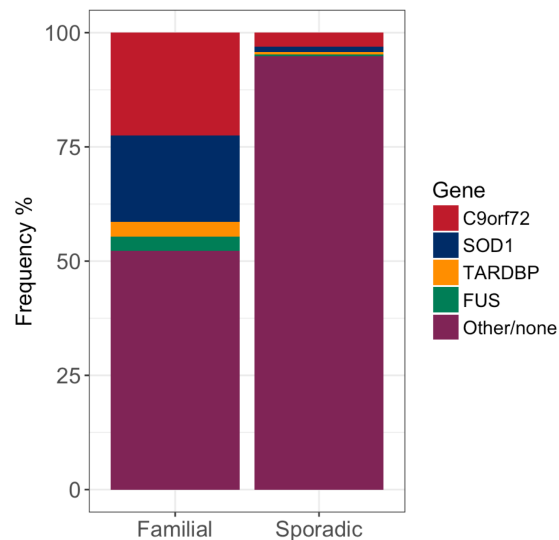


Figure 1.3: Frequency of major ALS genes in sporadic and familial ALS. Based on data from Zou et al. 2017.

inherited trait (Byrne et al. 2011). Recessively inherited cases are also described, though these are extremely rare (Maruyama et al. 2010). It has also become clear that mutations in the known ALS-associated genes are found in a small proportion of apparently sporadic cases, due to incomplete penetrance, early parental mortality, non-paternity or *de novo* mutation (Majounie et al. 2012; Chio et al. 2008).

An additional genetic contribution to the susceptibility to sporadic ALS is supported by twin concordance data, estimating heritability to be around 61% (Al-Chalabi et al. 2010). Several genes conferring increased risk of developing ALS have been identified, though the extent of genetic changes contributing to altered risk is not established.

Although over 30 candidate ALS genes have been identified, 48% of familial ALS cases (and 5% of sporadic cases) are attributable to mutations in four genes: *C9orf72*, *SOD1*, *TARDBP* and *FUS* (**Figure 1.3**; Zou et al. 2017).

C9orf72

Expansion of the hexanucleotide repeat (GGGGCC)_n in intron 1 of the *C9orf72* gene is the most common cause of familial ALS and FTD, found in 39% of familial and 7% of sporadic ALS cases in Europe and the USA, though is less common in Asia (Renton et al. 2011; Murray et al. 2011; Majounie et al. 2012; Zou et al. 2017). *C9orf72*-associated ALS and FTD are clinically and pathologically indistinguishable from sporadic ALS (Murray et al. 2011).

Three non-exclusive mechanisms by which the *C9orf72* hexanucleotide repeat expansion causes ALS and FTD are hypothesised, though the relative contribution of these mechanisms and the precise means of toxicity remain unresolved. Haploinsufficiency of the C9orf72 protein is observed in *post mortem* tissue, with reduced expression of isoform A (Murray et al. 2011). The C9orf72 protein is involved in autophagy and downregulation is hypothesised to increase the toxicity of aggregation-prone proteins, though evidence is conflicting (Sellier et al. 2016; Therrien et al. 2013; O’Rourke et al. 2016).

Sense and antisense transcription of the repeat region leads to the formation of RNA foci which are detectable in *post mortem* human brain and spinal cord and induced pluripotent stem cell (iPSC)-derived motor neurons (Murray et al. 2011; Lagier-Tourenne et al. 2013; Almeida et al. 2013). RNA foci interact with RNA binding proteins and might lead to widespread alterations in gene splicing, thereby mediating neurodegeneration (Cooper-Knock et al. 2014; Donnelly et al. 2013; Sareen et al. 2013; Prudencio et al. 2015).

Sense and antisense *C9orf72* repeat transcripts are also translated by a process known as repeat-associated non-ATG translation (RAN translation), also described in other repeat-associated neurodegenerative diseases, leading to the production non-canonical proteins from all reading frames (Cleary et al. 2013). The resultant dipeptide repeat proteins are neurotoxic and form p62-positive, TDP-43-negative

cytoplasmic and nuclear aggregates within neurons (Mori et al. 2013).

SOD1

The first identified ALS-associated genetic mutations were in the *SOD1* gene, encoding the antioxidant enzyme Superoxide dismutase [Cu-Zn] (SOD-1) (Rosen et al. 1993). Over 180 mutations in *SOD1* have been identified, accounting for 18.9% of familial ALS and 1.2% of sporadic cases in a recent meta-analysis, with some global variation (*SOD1* mutations being more common in Asian populations; Zou et al. 2017).

SOD1 ALS presents as typical ALS without cognitive impairment, otherwise indistinguishable from sporadic ALS, though bulbar onset is less frequent (Cudkovic et al. 1997; Yamashita et al. 2015).

TARDBP

Mutations in *TARDBP*, which encodes TDP-43, are associated with ALS (Neumann et al. 2006). *TARDBP* mutations account for an estimated 3.3% of familial and 0.5% of sporadic ALS cases (Zou et al. 2017). Although *TARDBP* mutations mostly manifest clinically as typical ALS (Sreedharan et al. 2008; Gitcho et al. 2008; Van Deerlin et al. 2008; Kirby et al. 2011), *TARDBP* mutations have also been identified in patients with FTD (Benajiba et al. 2009).

Over 50 ALS-associated mutations in *TARDBP* have been identified. Almost all of these fall within the C-terminal low complexity domain of TDP-43, a highly conserved region involved in protein binding and liquid-liquid phase separation in the formation of stress granules (Molliex et al. 2015).

FUS

Mutations in *FUS*, encoding RNA-binding protein FUS (FUS), account for an estimated 3.0% of familial and 0.4% of sporadic cases of ALS (Zou et al. 2017). *FUS* mutations usually cause typical ALS (Kwiatkowski et al. 2009), as well as juvenile-onset ALS and FTD (Baumer et al. 2010; Van Langenhove et al. 2010). In addition, *FUS* mutations have also been detected in patients with essential tremor (Merner et al. 2012).

FUS has both functional and structural features in common with TDP-43. Structural similarities include the presence of glycine-rich low complexity domains, an RNA recognition motif (RRM) and nuclear localisation sequence (NLS). Like TDP-43, FUS binds both DNA and RNA with a role in regulating gene expression and RNA splicing (Dormann et al. 2013). Unlike *TARDBP*, however, ALS-associated mutations in *FUS* have been identified throughout the length of the *FUS* gene, in particular the NLS and not confined to the low complexity domain as observed in *TARDBP* (Deng et al. 2014). Inclusions in *FUS*-ALS are immunopositive for FUS and not TDP-43 (Baumer et al. 2010).

Other Gene Mutations

Mutations in at least 28 other genes have been implicated in causing ALS, including genes encoding proteins involved in RNA processing (*HNRNPA1* and *MATR3*; Kim et al. 2013; Liu et al. 2016; Johnson et al. 2014; Leblond et al. 2016), axonal cytoskeletal and trafficking components (*NEFH*, *TUBA4A* and *DCTN1*; Al-Chalabi et al. 1999; Skvortsova et al. 2004; Smith et al. 2014; Puls et al. 2003; Munch et al. 2004; Munch et al. 2005), and proteins involved in autophagy and proteasomal pathways (*UBQLN2*, *SQSTM1* and *TBK1*; Deng et al. 2011; Fecto et al. 2011; Hirano et al. 2013; Le Ber et al. 2013; Teyssou et al. 2013; Cirulli et al. 2015; Freischmidt et al. 2015).

ALS as an oligogenic disease

Twin concordance data suggests that familial ALS does not capture the extent of the genetic contribution to ALS susceptibility (Al-Chalabi et al. 2010). A number of genes conferring increased risk of developing ALS have been identified, such as *NEK1* (Kenna et al. 2016), *ANG* (Es et al. 2011) and intermediate-length expansions in the *SCA2* CAG repeat (27-33 repeats associated with ALS, >33 mutations associated with spinocerebellar ataxia type 2 (SCA2); Elden et al. 2010), alongside genes that are associated with altered prognosis (for example *DAO* and *UNC13A*; Cirulli et al. 2015; Diekstra et al. 2012; Chio et al. 2013b).

Increasingly, ALS cases and families are being identified that harbour more than one ALS-associated gene mutation or variant, occurring more frequently than would be expected by chance (Blitterswijk et al. 2012; Bury et al. 2016). Oligogenic inheritance might go some way to explaining the mismatch between estimates of heritability and the reported incidence of familial cases; it may also indicate that current estimates of penetrance of well-characterised mutations such as the *C9orf72* hexanucleotide repeat expansion, largely based on kindreds potentially carrying multiple risk-modifying variants, are overestimated.

1.2.6 Pathomechanisms in ALS

No single hypothesis has emerged that satisfactorily explains the clinical and neuropathological changes of ALS. Epidemiological evidence suggests that ALS is a multi-step process, requiring the accumulation of cellular events that lead to neurodegeneration, somewhat akin to the accrual of somatic mutations in cancer (Al-Chalabi et al. 2014). The following section will discuss some of the leading hypothesised mechanisms in ALS pathogenesis.

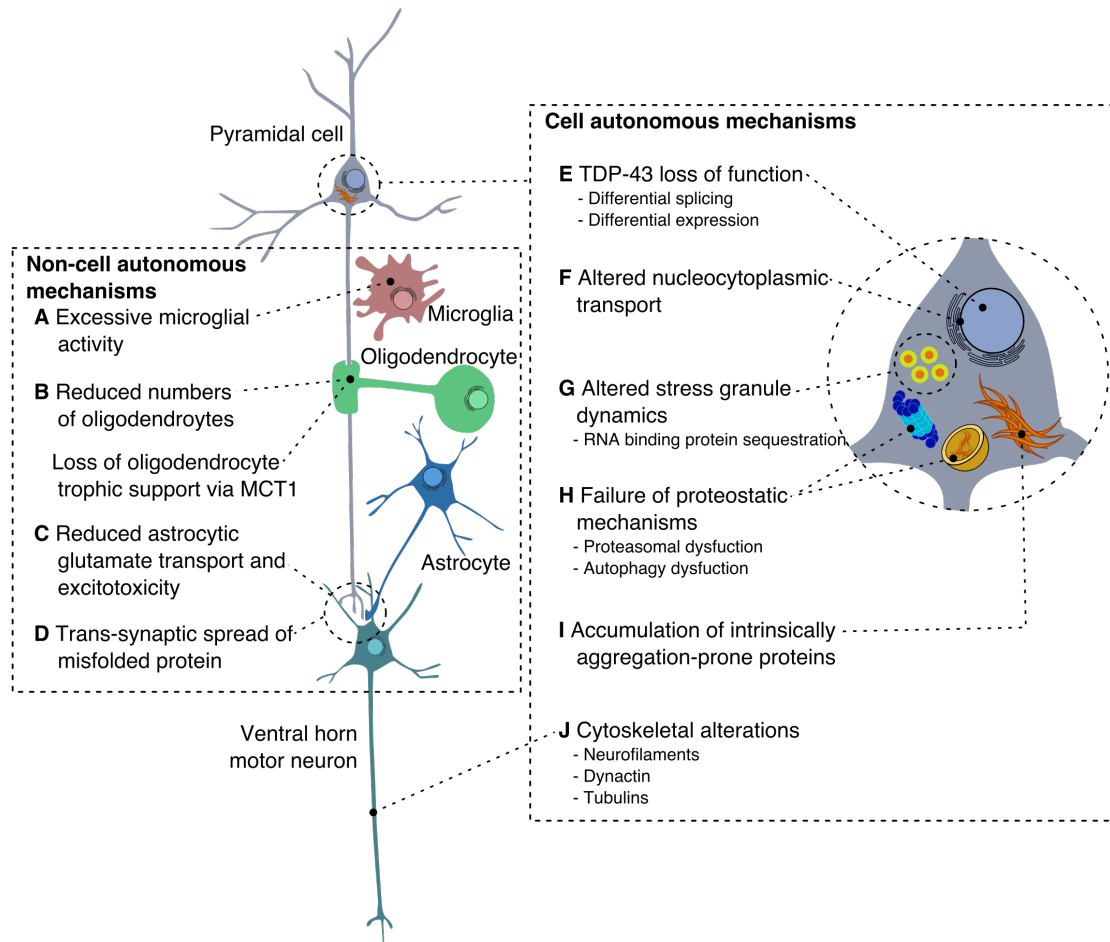


Figure 1.4: Proposed cellular mechanisms underlying ALS. Multiple cell-autonomous and non cell-autonomous mechanisms are proposed to contribute to ALS. **A** Microglial activity is observed in ALS and is an early feature in the SOD1^{G93A} mouse model. Downregulating microglial activity may have a modulating effect on ALS. **B** Oligodendrocyte numbers are reduced in ALS and loss of trophic support due to downregulation of monocarboxylic acid transporter 1 (MCT1) may contribute to neurodegeneration. **C** Alterations in astrocytic Excitatory amino acid transporter 2 (EAAT2) may increase extracellular glutamate and contribute to excitotoxic loss of motor neurons. **D** Trans-synaptic and intercellular transfer of misfolded proteins is a proposed mechanism of spread of neurodegeneration in ALS, occurring via EV-dependent and independent mechanisms. **E** Alterations in RNA processing including splicing and stability of certain pre-messenger RNAs (mRNAs) has been demonstrated due to TDP-43 loss of nuclear function. **F** ALS-associated mutations in the nuclear localisation sequence of FUS are associated with ALS and disruption of the nuclear localisation sequence of TDP-43 recapitulates TDP-43 aggregation in animal models. **G** proteins encoded by ALS-associated genes are incorporated into stress granules and ALS-associated mutations disrupt stress granule dynamics. This is proposed as a pre-aggregation stage. **H** Autophagy and ubiquitin-proteasome system (UPS) function decline with age. Mutations in UPS and autophagy genes disrupt protein degradation and are associated with ALS. Downregulation of UPS or autophagy leads to neurodegeneration in animal models. **I** TDP-43, FUS and other RNA-binding proteins encoded by ALS genes are highly aggregation prone, and demonstrate seeded aggregation *in vitro*. **J** Mutations in cytoskeletal proteins are associated with ALS.

Glial cells and non-autonomous mechanisms in ALS pathogenesis

Glial cells are of key importance to the health of CNS neurons, with functions including trophic support, innate immunity, synaptic remodelling and neurotransmitter metabolism (Jakel et al. 2017). Glial pathology in CNS regions affected by ALS have long been known, though their role in the disease whether causing, exacerbating or protecting against neurodegeneration has been hotly debated for decades.

Microglia Considered part of the innate immune system, microglia are myeloid cells derived from the monocyte-macrophage lineage (Jakel et al. 2017). Reactive microgliosis is a phenomenon observed in virtually all CNS pathologies including neurodegenerative disease, and ALS is no exception (McGeer et al. 1993); similarly, the controversy surrounding their role is not unique to ALS. In *post mortem* tissue, microglial pathology is found in the spinal cord ventral horn (sometimes in close contact with motor neurons) within the corticospinal tracts, subcortical white matter and motor cortex of patients with ALS (Kawamata et al. 1992; Brettschneider et al. 2012). Microglial morphology varies by region, with elongated and branched microglia found in the ventral horns compared with more rounded shape in the white matter of the corticospinal tract (Brettschneider et al. 2012).

Microglial activation has also been observed in the brain of ALS patients *in vivo* using positron emission tomography (PET) ligands for the peripheral benzodiazepine receptor, which is expressed by activated microglia (Turner et al. 2004; Zurcher et al. 2015). These studies have reproducibly demonstrated evidence of microglial activation in the precentral gyrus and brainstem of patients with ALS compared to healthy controls. This activation correlates with the burden of clinical UMN dysfunction.

An increase in microglial number is observed in the SOD1^{G93A} mouse model of ALS

and is detectable before the onset of motor phenotype, when neuronal loss is minimal and prior to astrocyte activation and T-cell infiltration (Alexianu et al. 2001; Saxena et al. 2009). Microglial morphology and phenotype appear to change through the disease course, with anti-inflammatory Interleukin (IL)-10-secreting microglia predominating in early disease and increased expression of pro-inflammatory cytokines IL-1 and Tumor necrosis factor alpha (TNF α) at later stages (Ohgomori et al. 2016; Gravel et al. 2016; Beers et al. 2011).

Evidence for a protective or detrimental role for microglia in SOD1 rodent models of ALS is somewhat contradictory. Inhibition of microglial activation and proliferation via blockade of the cytokine receptor Macrophage colony-stimulating factor 1 receptor (CSF1R) (Martinez-Muriana et al. 2016) and selective reduction of mutant SOD-1 in microglia (Boillee et al. 2006; Trias et al. 2016) have both been shown to improve survival, though ablation of proliferating microglia has no appreciable effect on survival (Gowing et al. 2008).

Little evidence of the role of microglia has been provided by TDP-43-based models. In the TDP-43 RLS8 mouse model, with motor neuron-restricted doxycycline-suppressible expression of a NLS-sequence deleted human TDP-43 (Walker et al. 2015), clearance of aggregated TDP-43 in motor neurons and recovery of motor function is dependent upon microglia, suggesting a possible protective role for microglia (Spiller et al. 2018).

Evidence from these rodent models must be interpreted with caution in the context of ALS, since neither overexpression of SOD-1 nor deletion of the TDP-43 NLS faithfully recapitulate the human disease. Nonetheless, the evidence for a detrimental effect of microglia in ALS has been deemed sufficient for clinical trials of a CSF1R inhibitor, Masitinib, to proceed in humans.

Astrocytes and excitotoxicity Widespread astrogliosis has been shown in spinal cord ventral horn, cerebral cortex and subcortical white matter of *post mortem* tissue from ALS patients (Schiffer et al. 1996; Nagy et al. 1994; Kushner et al. 1991). Reduced expression of the astrocytic glutamate transporter EAAT2 and reduced synaptosomal glutamate transport, essential for the maintenance of low extracellular glutamate concentrations, has also been described in *post mortem* specimens (Sasaki et al. 2000; Rothstein et al. 1995; Rothstein et al. 1992; Rothstein et al. 1996).

Spinal motor neurons appear to be particularly susceptible to toxicity from excessive glutamate, specifically via ionotropic glutamate receptor-mediated calcium influx (Van Den Bosch et al. 2000). This increased susceptibility may relate to lower expression of calcium-buffering proteins such as parvalbumin and calbindin when compared to motor neurons from regions infrequently affected by ALS, specifically the nuclei supplying the extraocular muscles and Onuf’s nucleus (Alexianu et al. 1994).

Elevated glutamate in CSF from ALS patients has been inconsistently demonstrated, suggesting that these reductions in EAAT2 may result in increased extracellular glutamate (Shaw et al. 1995). These findings, however, are not supported by *post mortem* studies which have instead found *reduced* levels of glutamate in spinal cord ventral horn (Tsai et al. 1991; Malessa et al. 1991); a more recent metabolomic study of CSF also found reduced levels of glutamate (Wuolikainen et al. 2011).

CSF from ALS patients is neurotoxic, in a motor neuron-preferential fashion; this effect can be abrogated by blocking ionotropic glutamate receptors, providing additional evidence that excitotoxicity is relevant *in vivo* (Sen et al. 2005). Interestingly, sensitivity to excitotoxicity may be exacerbated by *C9orf72* haploinsufficiency (Shi et al. 2018).

Oligodendroglia Reduced numbers of oligodendrocytes are also observed *post mortem* in motor cortex, corticospinal tract and ventral horn of patients with ALS accompanied by proliferation of NG2⁺ oligodendrocyte precursors (Kang et al. 2013). Oligodendrocytes provide myelination of CNS axons as well as trophic support for axons, including the provision of glycolytic substrates via MCT1. Loss of oligodendroglial MCT1 is neurotoxic and leads to an axonopathy in mice, and levels of MCT1 are decreased in the motor cortex of patients with ALS (Lee et al. 2012). This implies that oligodendrocyte loss might exacerbate neurodegeneration rather than occurring purely as a secondary consequence.

RNA biology in ALS

TDP-43 has roles in multiple aspects of RNA biology, including transcription regulation, splicing and micro-RNA (miRNA) processing (Lagier-Tourenne et al. 2010). The identification of mutations in genes encoding additional RNA binding proteins (*FUS*, *MATR3* and *HNRNPA1*) has further cemented RNA biology at the centre of efforts to understand the cellular mechanisms leading to neurodegeneration in ALS.

TDP-43 binds UG-rich motifs, mostly in intronic regions of pre-mRNA within the nucleus (Tollervey et al. 2011). The subcellular mislocalisation of TDP-43 observed in ALS *post mortem* tissue, whereby the normal nucleus-restricted TDP-43 immunoreactivity is lost and cytoplasmic TDP-43 observed instead, raises the possibility of a loss of nuclear function as a contributing mechanism in ALS (Neumann et al. 2006).

Knockdown of TDP-43 in mice leads to reduced expression of genes involved in synaptic activity and function, as well as leading to alternative splicing of a large number of genes (Polymenidou et al. 2011). Alternative splicing events have also been demonstrated in mouse models overexpressing wild-type and ALS-associated mutant human TDP-43, showing a dose-dependent effect of TDP-43 on splicing

(Arnold et al. 2013). Some evidence also suggests that alternative splicing events mediated by nuclear TDP-43 depletion may lead to the production of protein isoforms with increased fibrilization potential, implying that splicing dysregulation might exacerbate protein aggregation in ALS (Deshaies et al. 2018).

Differential splicing has also been demonstrated in LMNs from patients with ALS, affecting 24% of genes analysed, including 15 of the genes known to cause ALS and overlapping significantly with alternative splicing events detected in fibroblasts carrying *TARDBP* mutations (Highley et al. 2014). Gene ontology (GO) enrichment analysis of this data identified enrichment of genes involved in RNA splicing, protein localization and cytoskeletal organization in those differentially spliced or expressed.

Similar to TDP-43, FUS appears to stabilise long pre-mRNA (though there is only modest overlap between TDP-43 and FUS in terms of which pre-mRNAs they stabilise; Lagier-Tourenne et al. 2012). FUS also regulates pre-mRNA splicing, with differential splicing of genes involved in nervous system development and some overlap with alternative splicing events regulated by TDP-43 (Rogelj et al. 2012; Lagier-Tourenne et al. 2012).

Proteostatic mechanisms

Protein degradation mechanisms are of major interest across the neurodegenerative diseases, since their shared pathological hallmark is the accumulation of aggregated proteins. The two major cellular protein degradation mechanisms, the autophagy-lysosome system (responsible for bulk degradation of organelles and aggregates) and the UPS (responsible for degradation of damaged or misfolded proteins), deteriorate with age and may contribute to the pathogenesis of ALS (Saez et al. 2014; Vilchez et al. 2014; Keller et al. 2000; Lipinski et al. 2010).

The hallmark aggregates of ALS contain TDP-43 that has been polyubiquitinated suggesting that it has been targeted for degradation in the UPS or via autophagy

(Neumann et al. 2006). TDP-43 inclusions also co-localise with p62 (Sequestosome-1), an autophagy regulator that targets ubiquitinated proteins for autophagy and is mutated in a small proportion of ALS cases (Bjorkoy et al. 2005; Fecto et al. 2011).

Mutations in *VCP*, also seen in a minority of ALS patients as well as multisystem proteinopathy (MSP), are associated with defective maturation of autophagosomes (Johnson et al. 2010; Tresse et al. 2010). Inhibition of autophagy is sufficient to cause widespread neurodegeneration in mice, though this is not specific to the motor system (Hara et al. 2006).

Reductions in the levels and catalytic activity of proteasomal 20S β 5 subunit has been reported in ALS *post mortem* spinal cord (Kabashi et al. 2012). ALS-associated mutations in *UBQLN2* impair delivery of ubiquitinated proteins to the proteasome and overexpression of mutated forms of its product Ubiquilin-2 slows proteasomal degradation of proteins, suggesting *UBQLN2* mutations might cause ALS by impairing UPS function (Chang et al. 2015). Knock out of proteasomal subunits in the motor neurons of mice leads to a motor phenotype with recapitulation of some pathological features of ALS (Tashiro et al. 2012).

Protein aggregation and prion-like propagation

Protein aggregation is a feature common to neurodegenerative diseases. However, it remains unanswered whether aggregates are an upstream event leading to toxicity and cell death or reflect a cellular response to insult, for example sequestering toxic protein oligomers (Fang et al. 2014; Ojha et al. 2011).

TDP-43 is inherently aggregation prone *in vitro*, a tendency which is reliant upon C-terminal glycine-rich low-complexity domain that appears to be responsible for some protein-protein interactions (Johnson et al. 2009). Aggregation of TDP-43 is also accelerated by the presence of ALS disease-associated *TARDBP* mutations (Johnson et al. 2009). TDP-43 aggregation can be initiated in cells overexpressing

human TDP-43 by the presence of TDP-43 fibrils (Furukawa et al. 2011) or insoluble TDP-43 extracted from CNS tissue of patients with ALS or FTD (Nonaka et al. 2013; Smethurst et al. 2016). Overexpression of TDP-43 is also sufficient to cause TDP-43 aggregation (Liu-Yesucevitz et al. 2010).

Such behaviour is not restricted to TDP-43. Around 1% of the human genome encodes proteins that are predicted to have a high propensity to fibrilize (so-called prion propensity score), with a significant 12-fold enrichment of RNA-binding proteins (King et al. 2012). Of the 11 RNA-binding proteins with the highest prion propensity score, six have been associated with mutations that are causative of ALS (King et al. 2012).

Similarly, though not an RNA binding protein, ALS-associated mutations in *SOD1* lead to accelerated SOD-1 aggregation in cells, though no convincing relationship between *in vitro* aggregation propensity and age of ALS onset or survival in patients with different *SOD1* mutations has been demonstrated (Prudencio et al. 2009). *In vitro* intercellular seeding of misfolded mutant SOD-1 in conditioned medium has also been demonstrated and is inhibited by depletion of endogenous SOD-1 (using small interfering RNA (siRNA) knock-down in recipient cells) suggesting that this seeding proceeds by inducing misfolding in the endogenous wild-type protein (Grad et al. 2014).

The finding that aggregation can be initiated by "seeding" with preformed aggregates is of particular interest. This is because it provides a potential substrate for the observed spatial spread of ALS symptoms and signs through the neuraxis by a mixture of local intercellular means and more distant spread via axonal transport systems, the CSF or systemic circulation (Ravits et al. 2009).

Such a mechanism is widely accepted in prion diseases such as Creutzfeldt-Jakob disease (CJD), whereby misfolded prion protein (PRP^{Sc}) initiates misfolding of native prion protein (PRP^{C}) (Prusiner 1982). Compelling evidence of the spread of patho-

logical proteins has also been found in Parkinson's disease, where Lewy bodies have been observed within mesencephalic stem cell transplants, implying host-to-graft transmission of misfolded α -synuclein (Kordower et al. 2008).

Mechanisms of intercellular spread The means by which such misfolded proteins would spread between cells is debated. In CJD, spread through macropinocytosis (Wadia et al. 2008) and via tunnelling nanotubes (Gousset et al. 2009) has been suggested, though most evidence points to a major role for EVs (see **subsection 1.3.2** for further details of EV biology; Fevrier et al. 2004; Vella et al. 2007; Alais et al. 2008; Guo et al. 2016).

Similar mechanisms might lead to hypothesised "prion-like" spread in ALS, though work to establish the means by which spread might occur is limited and focuses mainly on EVs.

Excretion of wild-type human and mouse SOD-1 has been documented in exosomes (the most studied of the EV subtypes) purified using sucrose-density gradient ultracentrifugation (UC), though interestingly SOD-1^{G93A} may be released in smaller quantities (Gomes et al. 2007). Misfolded SOD-1 appears to localise to the outer surface of exosomes *in vitro* studied using immunogold labeled electron microscopy (EM), with some SOD-1 not localised to exosomes but free in conditioned medium (Grad et al. 2014).

Release of TDP-43 in EVs has been demonstrated in cell culture models overexpressing TDP-43, both on the surface and within the lumen of EVs (Nonaka et al. 2013; Feiler et al. 2015). Enrichment of TDP-43 has also been observed within the UC pellet of CSF from patients with ALS and FTD, again suggesting release within EVs (though the amount of EV TDP-43 did not differ between patients and controls; Feneberg et al. 2014). In N2a cells overexpressing TDP-43, exosomal secretion of TDP-43 is enhanced by autophagy or proteasomal inhibition, and elevated levels

of TDP-43 are found in exosomes extracted from *post mortem* brain tissue from patients with ALS compared to healthy controls (Iguchi et al. 2016). Furthermore, treatment with the inhibitor of exosome formation GW4869 also leads to an increase in insoluble TDP-43 in CNS tissue of mice expressing TDP-43^{A315T} (though not in TDP-43^{WT}) accompanied by exacerbation of behavioural phenotype. This implies that exosomal release of TDP-43 might in fact be protective (Iguchi et al. 2016).

One study using a luciferase reporter method to detect TDP-43 oligomers in combination with a microfluidic chamber system demonstrated intercellular spread of TDP-43, possibly indicative of transynaptic spread via EVs (Feiler et al. 2015). Alternative mechanisms of spread might also be important: rat neural stem cell-derived neurons and glial cells, transduced to produce wild-type TDP-43 and TDP-43 C-terminal fragments with concurrent proteasome inhibition, have been shown to produce cytoplasmic TDP-43 aggregates which, following death of cells containing aggregates, can be taken up by neighbouring neurons (Ishii et al. 2017). Although the mechanism of uptake in this model was not elucidated, it suggests that EVs are not the only possible means of spread of toxic protein species in ALS.

Although intercellular spread of pathogenic proteins is an attractive explanation for the observed spread of symptoms in ALS, evidence is largely based on *in vitro* over expression models with a lack of *in vivo* support and does not adequately explain some aspects of the disease, such as the presence of widespread fasciculations prior to onset of weakness that occurs in some cases.

Integrating disease mechanisms The means by which different mechanisms interrelate and contribute to ALS is not understood. As outlined in the multi-step hypothesis, it is likely that accrual of abnormalities in several pathways is required to cause ALS (Al-Chalabi et al. 2014). The heterogeneity of genes associated with ALS, however, suggest that defects in diverse pathways converge on the final common pathway of motor neuron degeneration and clinical ALS. Therefore, although

multiple steps are required to cause ALS, these steps are unlikely to be the same in all cases (Turner et al. 2013b; Van Damme et al. 2017). This concept has major implications for the success of potential treatments targeting specific disease mechanisms and suggests that some form of molecular classification of ALS may one day be a major factor in determining appropriate therapy for individual patients.

1.3 Biomarkers of ALS

The concept of a biological marker has been around since the mid 20th century, though the term biomarker first found use in clinical literature in the 1970s (Porter 1957; Order 1977). Perhaps the most widely accepted definition of a biomarker is that of the National Institute of Health Biomarkers Definitions Working Group, (2001), describing a biomarker as *a characteristic that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention.*

Biomarkers can be classified by their purpose, the mechanism of measurement (such as physical measurements, molecular tests, imaging markers), the pharmacological level at which they occur (molecular, cellular, tissue or organ) or the mechanism of action of an intended intervention (Aronson et al. 2017). The taxonomy of disease biomarkers by purpose has evolved, with recent classifications recognising six types of biomarker, being used to (1) predict risk of developing a disease, (2) diagnose disease, (3) measure the severity of disease, (4) measure the stage of disease, (5) predict or measure treatment response (theranostic) and (6) predict prognosis (Davis et al. 2015). Surrogate endpoints are a subset of biomarkers that are used in lieu of a clinical endpoint in the context of clinical trials (Biomarkers Definitions Working Group, 2001)

There have been over 100 trials of candidate therapies for ALS, with just two drugs

(riluzole and edaravone) resulting in positive outcomes (Talbot et al. 2018). These trials have generally relied on measuring variables that change slowly or occur infrequently. Typically, functional deterioration, decreasing muscle strength or differences in survival are compared between treated and untreated groups (Turner et al. 2013a). Consequently trials in ALS are long and costly, and expose potentially vulnerable and significantly disabled patients to the inconvenience of prolonged clinical trial involvement, including toxicity of (largely ineffective) drugs (Cudkowicz et al. 2010).

These issues might be partially mitigated by improvements in outcome measures to detect an early treatment effect (or lack thereof), occurring days or weeks after the commencement of treatment; biomarkers offer a way to accomplish this and enable early decisions on trial abandonment or continuation (Turner et al. 2015). Delayed enrollment due to the diagnostic latency of ALS and inappropriate stratification of patients are other reasons for negative trial outcomes that might benefit from the use of appropriate biomarkers, not to mention the clinical benefits of accelerated diagnosis of ALS (or its exclusion, in the case of mimicking conditions), improved prognostication and mechanistic insights that might be gained from ALS biomarkers.

Features that would constitute an ideal biomarker in ALS have been proposed (Turner et al. 2009). ALS biomarkers should be (1) sensitive and specific for the syndrome of ALS, preferably before the onset of symptoms, (2) able to reliably distinguish clinical phenotypes associated with prolonged survival at symptom onset, (3) predict regional involvement in advance of symptoms, (4) change in a predictable way with progression of disease to permit judgement of therapeutic response and (5) use easily accessible and affordable technology (Turner et al. 2009).

Additional desirable characteristics for biomarkers that have been established for other neurodegenerative diseases are also applicable in ALS, such as a link between biomarkers and neuropathological features and that biomarkers should be validated in neuropathologically confirmed cases (Lewczuk et al. 2018; Shaw et al. 2007).

Risk biomarkers of ALS

Currently, the only known biomarkers of risk of ALS are the presence of ALS-associated gene mutations, which are generally only sought in those already diagnosed with ALS or with a family history of ALS (Vajda et al. 2017). Notably, although the presence of an ALS-associated gene mutation might predict the lifetime risk of developing ALS for an individual, it cannot predict when disease onset might occur. Biomarkers that predict disease onset (for example by detecting cellular or neuronal network perturbations that predate symptoms) may well be useful in future to enable the timely initiation of treatment (Benatar et al. 2012)

Diagnostic biomarkers

The median latency from symptom onset to diagnosis in ALS is approximately 11 months, though much of this delay is attributable to a lack of recognition of the gravity of symptoms or appropriate referral in primary care (Cellura et al. 2012). A diagnostic test for ALS is unlikely, therefore, to greatly hasten diagnosis in most cases, though might be helpful in distinguishing ALS from disorders mimicking ALS (Turner et al. 2009).

Prognostic biomarkers

The heterogeneity of disease progression and survival in ALS is a major challenge in order to appropriately counsel patients with respect to their likely survival and progression of disability, as well as in the efficient conduct of clinical trials (Bowser et al. 2011). Efforts to overcome this heterogeneity using solely clinical parameters have produced models that, whilst accurate at predicting disease progression or survival at the group level (and might reduce the number of participants required for a clinical trial by as much as 20%), lack accuracy at the individual level (Kuffner

et al. 2015; Westeneng et al. 2018).

Some biochemical markers have been incorporated into datasets used to produce algorithms predicting disease progression, but the biofluid markers with the most convincing association with the rate of disease progression and survival (CSF neurofilament proteins) were not included in these analyses, and have not been included in survival prediction models (Kuffner et al. 2015). It is not known whether the addition of other biomarkers to these algorithms would improve the accuracy of predictions.

Although the sequential involvement of body regions in ALS is somewhat predictable, more accurate prediction of the pattern of regional involvement in an individual would also be useful in ALS by permitting better planning of symptomatic interventions (Turner et al. 2009).

Biomarkers of therapeutic efficacy

The prolonged nature of clinical trials in ALS might be circumvented with appropriate markers. Sensitive and rapidly-changing measures of disease activity would allow the early detection of treatment effect before differences in disability and survival are detectable (Turner et al. 2015).

Biomarkers that predict response to specific treatments are well established in other fields, particularly cancer, where the presence of specific gene mutations or receptor expression can predict response to certain therapeutic approaches (Legha et al. 1978; Sawyers 2004). Similar approaches might prove useful in ALS. At a simple level, a genetic test showing presence of an ALS-associated gene mutation might predict response to antisense oligonucleotide therapy (Miller et al. 2013). More broadly, biomarkers of specific aspects of ALS pathology, such as neuroinflammation or oxidative stress, might predict response to specific therapies in those with sporadic disease and allow selection of appropriate treatments.

Biomarkers of progression

Current measures of disease progression in ALS are unsatisfactory. The most widely employed measure is the revised ALS functional rating scale (ALSFRS-R), a 48-point functional scale which assesses bulbar, fine motor, gross motor and respiratory domains (Cedarbaum et al. 1999). The total ALSFRS-R declines in an approximately linear fashion and its level at presentation is associated with survival (Moore et al. 2004; Kaufmann et al. 2005; Proudfoot et al. 2016). It has several limitations, declining slowly over time with the possibility of increasing if a patient overcomes a physical disability (with the use of aids, for example); there is cross-over of domains and the validity of the total score as a measure of disease burden is questioned (Bakker et al. 2017).

Measurement of muscle strength is an attractive progression measure, since weakness is the primary symptom of ALS. Measurement of limb muscle strength may be achieved using manual muscle testing (MMT), summing or averaging muscle strength as measured by the Medical Research Council grade, or using more quantitative measures such as hand-held dynamometry (HHD) or maximum voluntary isometric contraction (MVIC). MMT, MVIC and HHD are comparable in terms of inter-rater and intra-rater reliability, but one study suggested that MVIC lacks sensitivity compared to MMT, most likely due to the smaller number of muscles tested, though MMT is itself a non-linear scale (Great Lakes A. L. S. Study Group, 2003). Furthermore, these approaches cannot measure strength in the bulbar territory limiting their usefulness as a global measure of disease burden.

Fluid biomarkers that are objective and sensitive to overall progression of ALS, independent of site of symptom onset, would therefore be a useful surrogate in therapeutic trials.

1.3.1 Cerebrospinal fluid biomarkers of ALS

CSF surrounds the brain and spinal cord filling the ventricles and subarachnoid space of the CNS (Sakka et al. 2011). Traditional models of CSF physiology purport that around 80% CSF is produced in the choroid plexus, a secretory epithelium found in the lateral, third and fourth ventricles, with 20% produced by the ependyma and parenchyma of the CNS (reviewed in Brinker et al. 2014). More recent studies suggest significant exchange of CSF and CNS interstitial fluid across the ependymal surface implying that, although bulk flow from the parenchyma may be low, the contribution of CNS interstitial fluid constituents to the non-water content of CSF is likely to be higher (Abbott 2004).

As well as proteins produced in the CNS, CSF contains abundant plasma proteins. These are found at lower concentrations than plasma, with their CSF concentration dependent upon concentration in blood, hydrodynamic radius, and the integrity of the blood-CSF barrier (Reiber 1980; Guldbrandsen et al. 2014; Begcevic et al. 2016). Although CSF is harder to obtain than blood or urine, necessitating lumbar puncture, its proximity to the site of pathology in ALS, the presence of CNS-specific proteins and the signal-to-noise ratio afforded by lower concentrations of abundant proteins when compared to blood, makes CSF the preferred biofluid for biomarker discovery in ALS.

A wide variety of CSF constituents have been studied in ALS, including neurotransmitters, aggregate components, markers of oxidative stress, amino acids, cytokines and axonal degeneration markers (Turner et al. 2009; Vu et al. 2017). More recently, experiments have taken a data-driven approach utilising technologies such as mass spectrometry and nuclear magnetic resonance (NMR) to perform proteomic and metabolomic analyses for biomarker discovery (Blasco et al. 2013; Barschke et al. 2017). Some of the major groups of ALS biomarkers are outlined below.

Neurofilament proteins The most well-characterised CSF biomarkers of ALS currently are the neurofilament proteins, intermediate filaments that constitute the major cytoskeletal component of the axon. They are composed of heteropolymers of neurofilament light, medium and heavy chains (Wang et al. 2012). Elevated CSF and blood neurofilament levels are not only seen in ALS: neurofilament levels rise following a diverse range of insults leading to axonal damage from head trauma to multiple sclerosis (Shahim et al. 2017; Hakansson et al. 2017).

Although CSF neurofilament levels are accurate in distinguishing ALS patients from healthy controls, with sensitivity and specificity in excess of 90% in more recent studies, performance differentiating ALS from other neurological diseases (and specifically ALS mimics) is less well-established (Brettschneider et al. 2006; Zetterberg et al. 2007; Lu et al. 2015; Ganesalingam et al. 2011). The range of reported sensitivity to differentiate ALS from mimics is 78-93% for pNFH, 77-89% for neurofilament light chain (NFL); and specificity 63-94% for pNFH and 63-89% for NFL (Steinacker et al. 2016; Feneberg et al. 2018; Poesen et al. 2017; De Schaepdryver et al. 2017; Rossi et al. 2018).

Neurofilament levels have been widely found to correlate with the rate of disease progression and are associated with shortened survival (Zetterberg et al. 2007; Boylan et al. 2009; Ganesalingam et al. 2011; Boylan et al. 2013; Lu et al. 2015; Steinacker et al. 2016; McCombe et al. 2015; Poesen et al. 2017; De Schaepdryver et al. 2017; Rossi et al. 2018). Though few studies have analysed longitudinal variation, levels appear to be stable over time and are lower in those taking riluzole, making neurofilaments an attractive measure of disease activity in the context of clinical trials (Lu et al. 2015; McCombe et al. 2015). However, changes in neurofilament levels following the initiation of riluzole have not been demonstrated.

Neurofilament levels have been measured in CSF and blood of asymptomatic subjects carrying ALS-associated genetic mutations, in whom no elevation of levels were detected (Weydt et al. 2016). More recent analysis suggests that a rise in NFL levels

is detectable in CSF in the 12 months prior to symptom onset in this group, making neurofilaments a candidate for detection of disease onset prior to the development of clinical symptoms (Benatar et al. 2018).

There has also been significant interest in CSF levels of the microtubule associated protein Tau as a potential diagnostic biomarker of ALS. Several studies have explored levels of phosphorylated Tau and total Tau as well as the ratio of phosphorylated-to-total Tau (Grossman et al. 2014; Wilke et al. 2015; Scarafino et al. 2018). The results of these studies have varied widely, suggesting at present that measurement of Tau protein is of limited value in the diagnosis of ALS.

Aggregate components Naturally, target driven studies have focused on protein components of aggregates observed in ALS, in particular TDP-43, SOD-1 and Cystatin-C. Although TDP-43 has been shown to be elevated in CSF of ALS patients in several studies in comparison to healthy and disease controls this has not yielded a useful diagnostic test (Steinacker et al. 2008; Kasai et al. 2009; Noto et al. 2011; Feneberg et al. 2014). Counterintuitively, decreased levels of TDP-43 were associated with shortened survival in one study (Noto et al. 2011). Most CSF TDP-43 may also be blood-derived, suggesting that its measurement is unlikely to reflect the underlying pathological burden (Feneberg et al. 2014).

Although different banding patterns are observed in *SOD1* mutation carriers on immunoblot, overall SOD-1 activity and levels (including the levels of misfolded SOD-1) are not different in patients with ALS due to *SOD1* mutations compared with ALS patients without *SOD1* mutations or controls (Jacobsson et al. 2001; Zetterstrom et al. 2011). Again, SOD-1 may prove to be useful as a pharmacodynamic marker in antisense oligonucleotide trials for *SOD1*-ALS (Winer et al. 2013; Miller et al. 2013).

Cystatin-C, a highly-abundant CSF protein and component of Bunina bodies, a

type of protein inclusion found in LMNs in ALS, has been shown to be decreased in the CSF of patients with ALS in both candidate-driven and proteomic studies (Ranganathan et al. 2005; Tsuji-Akimoto et al. 2009; Ryberg et al. 2010; Wilson et al. 2010; Collins et al. 2015). Diagnostic accuracy appears poor, however, with sensitivity of 52% at best with corresponding specificity of 52% (differentiating ALS from all controls, 89% differentiating ALS from mimics); lower levels of Cystatin-C are also associated with worse survival in ALS patients (Wilson et al. 2010).

Cytokines and growth factors A number of cytokines and growth factors have been explored in the CSF of ALS, with several studies using a multiplex immunoassay to measure up to 33 cytokines simultaneously in CSF, serum or plasma from patients with ALS, other neurological diseases and healthy controls (Mitchell et al. 2009; Tateishi et al. 2010; Furukawa et al. 2015; Tanaka et al. 2006; Guo et al. 2017). Of the markers studied, Granulocyte colony-stimulating factor (CSF3) has been consistently elevated in all five studies examining CSF, IL-17 has been elevated in four of these five and Monocyte chemoattractant protein-1 (MCP1) and Vascular endothelial growth factor (VEGF) in three of five CSF studies. Elevation of IL-17 (along with IL-23, an activator of T-helper (Th)17 cells promoting chronic tissue inflammation) was demonstrated in a further study (Rentzos et al. 2010).

Diagnostic accuracy of CSF cytokines has been little studied. One study reported a sensitivity of 88% and specificity of 91% for a panel of five cytokines (Granulocyte-macrophage colony-stimulating factor (CSF2), IL-2, IL-6, IL-10 and IL-15) selected from a panel of 30, though this lacks subsequent validation (Mitchell et al. 2009).

Several cytokines have been associated with clinical parameters. In one study, higher CSF levels of Fibroblast growth factor 2, also known as FGF basic (FGF2) and CSF and serum VEGF were associated with improved survival, whilst CSF Interferon gamma (IFNG) was associated with worsened survival in Cox proportional hazards modelling (Guo et al. 2017). In the same study, MCP1 and IFNG in CSF and serum

correlated positively with disease progression rate (PR), and VEGF, FGF2 and C-C motif chemokine 3, also known as monocyte inflammatory protein 1 α (CCL3) in CSF and serum were negatively correlated with PR.

The CSF proteome in ALS

Various studies have utilised proteomic analysis of CSF to identify novel ALS biomarkers over the last 14 years, encompassing a range of mass spectrometry techniques accompanied by diverse methods to try to increase the relevance, depth and quantitative accuracy the identified proteome. Although several candidates have overlapped between studies, biomarkers identified by proteomics have often failed independent validation.

Candidate protein biomarkers in ALS identified in two or more studies are summarised in **Table 1.1**.

	Ranganathan et al. 2005	Pasinetti et al. 2006	Ranganathan et al. 2007	Brettschneider et al. 2008	Ryberg et al. 2010	Neuhoff et al. 2012	Varghese et al. 2013	Collins et al. 2015	Chen et al. 2016b
Cystatin-C	↓	↓	↓	-	↓	-	-	↓	-
Transthyretin	↓	-	↓	-	↑	-	-	-	-
Zinc-alpha-2- glycoprotein	-	-	-	↑	-	-	-	-	↓
Beta-2 microglobulin	-	-	-	↓	↓	-	-	-	-
Alpha-1-antitrypsin precursor	-	-	-	↑	-	↑	-	-	↓
Ribonuclease pancreatic	-	-	-	-	↓	-	-	-	↓
Hemoglobin subunit alpha	-	-	-	-	↑	-	-	-	↑
Chitotriosidase-1*	-	-	-	-	-	-	↑	-	-
Chitinase-3-like protein 2	-	-	-	-	-	-	↑	↑	-

Table 1.1: Candidate protein biomarkers identified in two or more proteomic studies comparing CSF of ALS patients with healthy or disease controls. *Chitotriosidase-1 has been separately validated in a study using ELISA. ALS - amyotrophic lateral sclerosis; CSF - cerebrospinal fluid.

The first proteomic study of ALS CSF, published in 2004 (Ramstrom et al. 2004), used liquid chromatography Fourier transform ion cyclotron resonance (FTICR) mass spectrometry of trypsin-digested CSF to derive mass spectrograms from 8 healthy control and 9 ALS subjects. These were then used to select features for classification of a test set comprising 2 healthy control and 3 ALS samples, of which 2 control and 2 ALS samples were classified correctly. Although no candidate markers were identified in this analysis and evaluation of classifier performance was limited, it was the first mass spectrometric analysis of CSF from ALS patients and suggested the possibility of classifiers based on multiple CSF protein features.

Three subsequent studies used surface-enhanced laser desorption-ionisation-time of flight (SELDI-TOF) mass spectrometry analysis for top-down proteomics of ALS CSF (Ranganathan et al. 2005; Pasinetti et al. 2006; Ryberg et al. 2010). The first of these studies created a panel of biomarkers based on differentially abundant mass peaks in the mass spectra of 15 ALS and 21 healthy control CSF samples. Of 366 mass peaks, identified using two SELDI-TOF chips, a total of 52 peaks with significant differences in abundance at the $p < 0.05$ level (not corrected for multiple comparisons) were identified (Ranganathan et al. 2005). A machine learning algorithm was also used to identify features differentiating ALS and control CSF, producing several peak panels that were validated in two separate test groups (10 control, 8 ALS and 7 control, 5 ALS), including 10 neurological disease control subjects. The best-performing panel of 19 peaks had sensitivity 80% and specificity 100% in the limited test set and a smaller panel of 3 peaks demonstrated sensitivity 80% and specificity of 71% in the same test set.

Subsequent characterisation of the proteins responsible for these 3 peaks identified neurendocrine protein 7B2, Cystatin-C and transthyretin. Decreases in Cystatin-C and Transthyretin were subsequently validated by immunoblot of CSF in a cohort overlapping with that used in proteomic analysis and immunohistochemistry of spinal cord.

In the second study, CSF from 36 ALS patients and 21 healthy controls was subjected to SELDI-TOF mass spectrometry, which identified 30 peaks, with reduced intensity observed in 3 peaks. These 3 peaks in combination provided sensitivity of 91% and specificity of 97% distinguishing between control and ALS CSF within sample (i.e. no test set was used; Pasinetti et al. 2006). Reductions in the peak intensity of the 3 proteins were validated using a further set of CSF samples from patients with ALS, healthy controls and patients with peripheral neuropathy and Parkinson's disease. Two of the 3 proteins were also subsequently purified and identified as Cystatin C (in agreement with Ranganathan et al. 2005, and further validated using ELISA) and a fragment of Neurosecretory protein VGF.

The most recent SELDI-TOF mass spectrometry study used individual samples from 100 ALS patients, 100 disease controls and 41 healthy controls (Ryberg et al. 2010). This identified 61 peaks with differential abundance between ALS and controls. Again, a univariate approach was used to generate candidate biomarkers, identifying elevation of a 23 kDa peak, corresponding to CRP, validated using ELISA; reduction in Cystatin-C at 13.38 kDa; and reduction in native Transthyretin alongside elevation of modified Transthyretin species. Multivariate analysis using machine learning was also used to construct a classifier incorporating 41 mass peaks, though this only resulted in modest estimated performance (sensitivity 63% and specificity 94%) on 10-fold cross-validation.

Two studies in 2008 and 2013 used two-dimensional gel electrophoresis to identify differentially abundant protein in CSF of ALS patients and healthy control subjects, both using pooled CSF samples (Brettschneider et al. 2008; Mendonca et al. 2012). Differentially abundant spots were then tryptically digested and the proteins identified using matrix-assisted laser desorption-ionisation-time of flight (MALDI-TOF) mass spectrometry. The first study employed peptide mass fingerprinting, identifying upregulation of two proteins (Alpha-1-antitrypsin precursor and Zn-alpha-2-glycoprotein) and downregulation of three (Caeruloplasmin precursor,

Beta-2-microglobulin and a third protein that was not identified; Brettschneider et al. 2008). Subsequent within-sample validation was performed for Zn-alpha-2-glycoprotein, confirming upregulation of this protein.

The later study employed tandem mass spectrometry of tryptically-digested differentially abundant spots, identifying 17 proteins in ALS that were not found in healthy control CSF (Mendonca et al. 2012).

A further study published in 2012 used 3 mass spectrometric techniques to further explore the CSF proteome in ALS (Neuhoff et al. 2012). Individual CSF samples from 35 ALS patients and 23 neurological disease controls were separated using weak cation exchange chromatography and analysed using MALDI-TOF mass spectrometry. A random forest predictor was then used to rank peaks and create a classifier based on a subset of MALDI-TOF spectra, resulting in a cross-validated sensitivity of 80.7% and specificity of 78.3%. Selected peaks were then identified using liquid chromatography-tandem mass spectrometry (LC-MS/MS) and validated in a separate cohort using a combination of ELISA and targeted mass spectrometry. Interestingly, although multivariate classifier performance was disappointing, several of the identified proteins overlapped with previous studies (Alpha-1-antitrypsin, Cystatin-C and Neurosecretory protein VGF) as well as two novel candidates (Osteopontin and Chromogranin A).

More recent proteomic analyses have used LC-MS/MS methodologies to identify biomarker candidates, using a mixture of techniques such as abundant protein depletion, prefractionation and isobaric labelling to increase proteomic depth and accuracy of quantification.

A 2013 study used pooled CSF from 10 ALS and 10 healthy control subjects which was subjected to immunoaffinity depletion of abundant proteins, labelled with isobaric tags and prefractionated using strong cation exchange chromatography prior to LC-MS/MS using an orbitrap tandem mass spectrometer (Varghese et al. 2013).

This yielded 819 protein identifications; candidates were selected purely on the basis of fold change between control and patient pools and included elevation of CHIT1, Osteopontin, CHI3L1 and CHI3L2 in ALS. Subsequent ELISA validation in a separate cohort supported a marked elevation of CSF CHIT1, accompanied by increased CHIT1 enzymatic activity, and elevated CHI3L2 and Osteopontin levels.

Two subsequent publications by different groups have used ELISA to measure CHIT1 in CSF, concurring with this proteomic analysis (Chen et al. 2016a; Steinacker et al. 2018). One study also demonstrated correlation of CHIT1 with disease progression rate and expression of CHIT1 in the corticospinal tract of patients with ALS (Steinacker et al. 2018).

A substantial improvement in proteomic depth was achieved in a 2015 LC-MS/MS-based study, using immunoaffinity-depleted CSF from 225 subjects, including 90 patients with ALS, combined into 25 pools (Collins et al. 2015). Samples were quantified using MS2-level label free quantification (spectral counting) on an orbitrap instrument without prefractionation, identifying 1712 unique proteins with at least two unique peptides.

Univariate analysis yielded 123 proteins differentially expressed with false discovery rate (FDR)-adjusted p -value < 0.05 . GO enrichment analysis suggested that these proteins were associated with extracellular matrix organisation, cell morphogenesis and differentiation, inflammatory responses and sensory nerve fasciculation. In-cohort validation was performed in CSF for three differentially abundant proteins using ELISA (Cystatin-C and Complement C3) and immunoblot (Secretogranin 1). Staining of *post mortem* spinal cord was performed for two identified candidate markers: the extracellular matrix component Tenascin R, which demonstrated reduced immunofluorescence, and Eukaryotic translation initiation factor 4E, a component of stress granules, that was observed in cytoplasmic inclusions and nuclear granules in ALS motor neurons, but not in healthy control motor neurons.

Four proteins were also selected to produce a support vector machine classifier, which was externally validated using 25 samples individually characterised by LC-MS/MS (18 ALS, 4 healthy control and 3 pooled other neurological disease samples), yielding area under the curve (AUC) of 0.94, sensitivity of 83% and specificity of 100%. Differences in the abundance of the four proteins, WD repeat-containing protein 63, SPARC-like protein 1, Amyloid-like protein 1 and Cell adhesion molecule 3, were confirmed in individual samples taken from the samples used in the training set.

Despite the impressive proteomic depth compared to previous studies, this analysis did not demonstrate increased abundance of CHIT1 in ALS as observed in Varghese et al. 2013, though elevation in CHI3L2 was observed. This may relate to the well-documented off-target effects of abundant protein depletion (the two experiments used different immunoaffinity depletion columns; Gunther et al. 2015), or that CHIT1 levels in the neurological disease control samples negated any elevation in ALS compared to control samples in the later study.

The most recent study using LC-MS/MS again employed immuno-depletion of abundant proteins, isobaric tag labelling and two-dimensional chromatographic separation (strong cation exchange and reverse-phase liquid chromatography) with a hybrid quadrupole time-of-flight mass spectrometer used to identify and quantify proteins in pooled CSF from patients with ALS, neurological disease and healthy control subjects (Chen et al. 2016b). The proteomic depth achieved was limited, with only 248 proteins identified. Since only one pool of ALS or control CSF was compared, isobaric tag ratios >1.2 or <0.8 were used to denote upregulated or downregulated expression, with 14 proteins elevated in ALS samples and 21 decreased by these criteria, including a number of highly abundant CSF proteins. Three candidates - Glutamate receptor 4, Leucine-rich alpha-2-glycoprotein and Insulin-like growth factor II - were selected and underwent validation by immunoblot, demonstrating elevation of Glutamate receptor 4 and a decrease in Insulin-like growth factor II in subjects with ALS within the pooled cohort.

Overall, proteomic studies of ALS CSF have produced several promising biomarker candidates, though with the exception of CHIT1, attempts at validation have generally proven them to be of limited use. No proteomic biomarkers yet match the effectiveness of neurofilaments, which remain the most promising ALS biomarkers.

1.3.2 Extracellular vesicles

EVs are lipid bilayer-bound extracellular structures that carry a cargo protein, as well as nucleic acids (principally RNA species, but also deoxyribonucleic acid (DNA); Bellingham et al. 2012; Huang et al. 2013; Thakur et al. 2014; Thompson et al. 2016).

Most work relates to two EV subtypes, exosomes and microvesicles (apoptotic bodies are also sometimes considered a subtype of EV), classified according to their cellular origin. Exosomes are derived from the endosomal compartment by invagination of the endosomal membrane to form multivesicular bodies (MVBs) containing multiple intraluminal vesicles (ILVs) (Colombo et al. 2014). MVB formation proceeds by mechanisms utilising the endosomal sorting complex required for transport (ESCRT) complex and via ESCRT-independent mechanisms such as those involving ceramides (Trajkovic et al. 2008). The MVB is then transported to and fuses with the plasma membrane, leading to release of exosomes into the extracellular milieu (**Figure 1.5**; Heijnen et al. 1999). Microvesicles form directly by outward blebbing of the plasma membrane, via multiple mechanisms including those dependent upon ESCRT proteins (Muralidharan-Chari et al. 2009).

Exosomes and microvesicles have similar properties and cargo, though microvesicles tend to be larger in size (50-2000 nm diameter compared to 40-150 nm for exosomes) and have a more diverse morphology (Thompson et al. 2016). EV subtypes can also be distinguished on the basis of their protein cargo, with smaller EVs, corresponding to exosomes, expressing higher levels of Tumour susceptibility gene 101 protein (TSG101), Syntenin-1 (SDCBP), ADAM10 and Annexin XI; and larger EVs

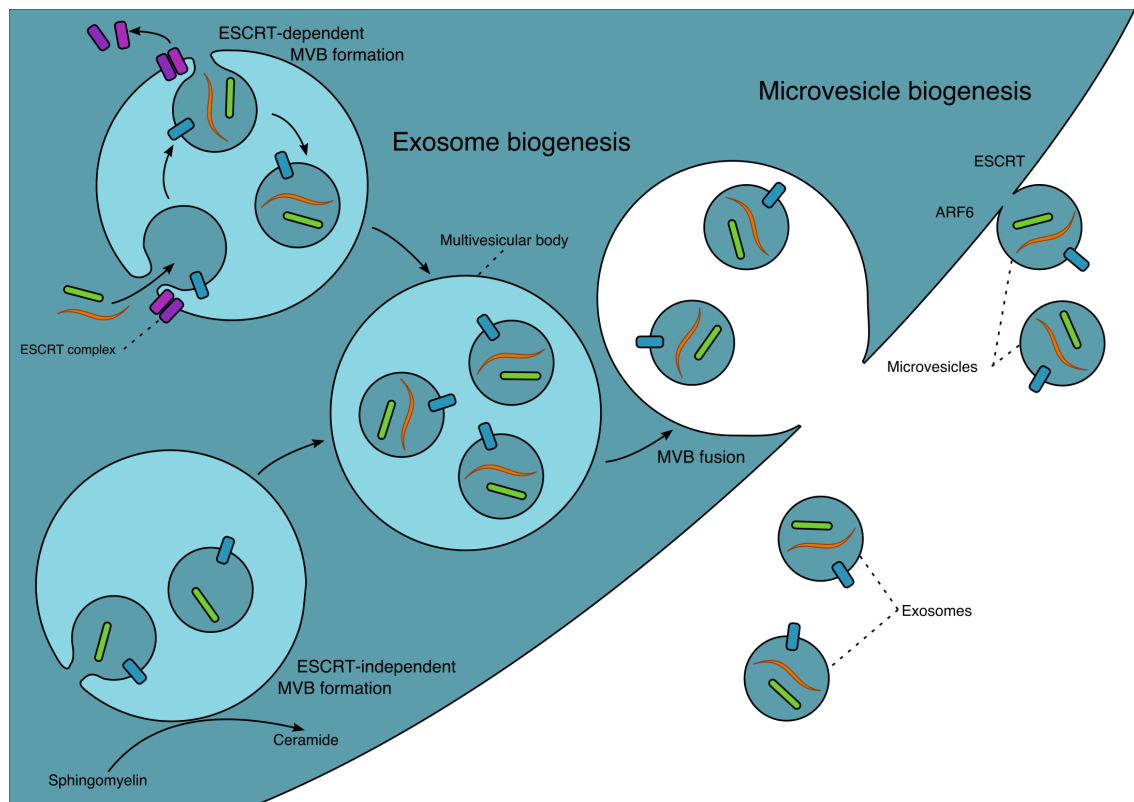


Figure 1.5: Biogenesis of the main EV subtypes, exosomes and microvesicles. Exosomes form by inward blebbing of the endosomal limiting membrane, via ESCRT-dependent and ESCRT-independent mechanisms (the latter involving ceramides). The thus produced MVB is then transported to then fuses with the plasma membrane, leading to release of exosomes. Microvesicles form directly by outward blebbing of the plasma membrane via mechanisms involving ESCRT components and ADP-ribosylation factor 6 (ARF6). EV - extracellular vesicle; ESCRT - endosomal sorting complex required for transport; MVB - multivesicular body; ARF6 - ADP-ribosylation factor 6. Adapted from Thompson et al. 2016.

expressing higher levels of Actinin-4, Mitofilin and Lysosome-associated membrane glycoprotein 2 (LAMP2) (Kowal et al. 2016).

EVs are produced by neurons and glia of the CNS and are involved in many key biological processes within the CNS including synaptic plasticity and pruning, trophic support of neurons and maintenance of the myelin sheath and intercellular signalling (Thompson et al. 2016). EVs have been implicated in several neurodegenerative diseases, primarily as vectors for intercellular spread of misfolded proteins (see **subsection 1.2.6**). CSF EVs contain proteins characteristic of neuronal and glial cell types, indicating that CSF EVs have, at least in part, CNS origin (Chiasserini et al. 2014). Proteins implicated in neurodegenerative diseases, including tau, α -synuclein, TDP-43 and amyloid precursor protein, have also been identified in CSF EVs (Saman et al. 2012; Chiasserini et al. 2014; Feneberg et al. 2014; Stuendl et al. 2016).

The intracellular origin of EVs makes them an attractive source of potential biomarkers of neurodegenerative diseases that might reflect underlying pathological processes. However, significant challenges impede the development of CSF EV-based biomarkers, particularly the paucity of CSF EVs (approximately 1000 fold lower than the concentration in serum) and the inefficiency of current isolation techniques (Thompson et al. 2016).

EV protein cargo has been studied for its biomarker potential in neurodegenerative diseases using several different approaches. A proteomic study using serum EVs purified using UC, identified decreased expression of SDCBP in EVs from patients with Parkinson's disease compared to controls (Tomlinson et al. 2015). Target-driven approaches have identified changes in phosphorylated tau isoforms in CSF EVs at different stages of Alzheimer's disease (Saman et al. 2012).

EV surface expression of the neuronal lineage marker L-1 cell adhesion molecule (L1CAM), presumed to denote CNS origin of EVs and expressing higher levels of Neuron-specific enolase and NFL, has been used to perform immunocapture of blood

EVs enriched in neuronal proteins (Mustapic et al. 2017). Elevation in the level of α -synuclein in EVs thus extracted has been demonstrated in Parkinson's disease patients (Shi et al. 2014; Fiandaca et al. 2015). Alterations in the levels of Alzheimer's disease-associated proteins and decreasing levels of synaptic proteins have also been identified in L1CAM-expressing plasma exosomes of patients with Alzheimer's disease, some of which are present at presymptomatic timepoints (Goetzl et al. 2015; Goetzl et al. 2018).

A single study has explored CSF EV protein in ALS (Feneberg et al. 2014). This study used a targeted proteomic approach to measure TDP-43 levels in UC-extracted CSF EVs from patients with ALS, FTD and healthy control subjects, demonstrating no difference in the levels between groups.

This thesis aims to use LC-MS/MS-based proteomic analysis of CSF and CSF-derived EVs to develop novel biomarkers of ALS.

1.4 Aims of the work presented in this thesis

1. Investigate the whole CSF proteome in ALS to identify novel candidate biomarkers of ALS (Chapter 3)
2. Validate the identified candidate biomarkers and explore their association with clinical and biochemical parameters (Chapter 4)
3. Establish a method of CSF EV isolation and compare this with current standard methods (Chapter 5)
4. Elucidate the origin of CSF EVs and their relevance to neurodegenerative diseases (Chapter 5)
5. Examine the CSF EV proteome in ALS as a source of novel candidate biomarkers (Chapter 6)

2. Materials and Methods

The experimental chapters of this thesis use LC-MS/MS proteomics to identify candidate biomarkers of ALS in whole CSF and CSF-extracted EVs. This methods chapter describes details of sample preparation, data acquisition and data analysis pertinent to biomarker discovery experiments using mass spectrometry proteomic analysis.

The principles of size exclusion chromatography, used in **Chapters 5** and **6** as a means of EV extraction, are also described. Finally, details of experimental methods applicable across the results chapters, including participant recruitment and characterisation, collection, handling and storage of samples, and statistical analysis methods will also be outlined.

2.1 Proteomic analysis

Proteins are the primary effector molecules of all known life, performing and controlling almost all biological processes across all organisms (Aebersold et al. 2016). The proteome describes all expressed proteins in a given system, comprising a complex network of interacting molecules with diverse function, structure and biochemistry. Proteomic analysis not only allows the simultaneous quantification of multiple analytes, but enables the integrated analysis of this network to elucidate pathways perturbed in an experimental paradigm or disease state (Bensimon et al. 2012; Zhu et al. 2001).

Mass spectrometry (MS) methods are the mainstay of modern proteomic analysis (though antibody and aptamer-based proteomic methods are also available; Gold et al. 2010). The central principal of MS-based proteomics is the high-precision

measurement of the mass-to-charge ratio (m/z) of ionised proteins, proteolytically cleaved peptides and their fragments in the gas phase. Most engineering and bioinformatic solutions in proteomics have focused on bottom-up proteomics, which describes experiments in which proteins are proteolytically cleaved prior to MS rather than the intact protein analysis of top-down proteomics (Aebersold et al. 2016). The resulting MS data from bottom-up proteomics can then be used to infer the protein composition of samples either by the presence of proteolytic peptides unique to a particular protein, as in tandem mass spectrometry (MS/MS), or the resulting pattern of peptides (peptide mass fingerprinting).

2.1.1 Proteomic analysis in biomarker development

Simultaneously measuring hundreds or even thousands of proteins (so-called shotgun proteomics), including post-translational modifications (PTMs) and isoform data, is a naturally attractive means of biomarker discovery, theoretically circumventing a great deal of effort, expense and sample use when compared to multiple candidate-driven studies.

However, proteomic analysis of biological fluids is not without limitations. The complexity of body fluids, particularly plasma and serum, is such that highly abundant proteins impede the identification of medium or low-abundance proteins, reducing the dynamic range and necessitating sample prefractionation or abundant protein depletion, which in turn increase preanalytical variability (Parker et al. 2014; Keshishian et al. 2015). The measured abundance of proteins in shotgun proteomics is generally only reported relative to other samples, rather than absolute values (Zhang et al. 2013). Furthermore, translation of candidate biomarkers identified in shotgun proteomic experiments usually requires additional validation using targeted proteomics or immunoassays, increasing costs (Anderson 2012).

Nonetheless, biofluids have been subjected to proteomic analysis to identify biomark-

ers in the research of a broad range of diseases. Earlier proteomic biomarker development studies predominantly featured the time-of-flight (TOF) MS technologies SELDI-TOF and MALDI-TOF. In MALDI-TOF, samples are co-crystallised with a proton-donating matrix onto a metal surface and ultraviolet laser light is used to induce desorption and ionisation of the sample. Ions are then accelerated by an electric field of known strength and the time taken for ions to travel a known distance before colliding with a detector is used to calculate the m/z . SELDI-TOF is based on similar principals, though the metal surface of MALDI-TOF is replaced by an array of planar chromatographics that allow separation of sample constituents by their chemistry, thereby reducing complexity (Liu 2011; Hutchens 1993).

More recently MS/MS methodologies, particularly LC-MS/MS have gained popularity and form the basis of the analytical methods used in this thesis.

2.1.2 LC-MS/MS

LC-MS/MS is commonly used for bottom-up proteomics and combines separation of peptides using liquid chromatography with ionization in-line, measurement of the m/z of precursor peptide ions (MS1) and their fragmentation spectra (MS2) to identify the protein constituents of a mixture and provide quantitative information. LC-MS/MS is applied to data-dependent acquisition (DDA) and data-independent acquisition (DIA) discovery proteomic and targeted proteomic applications using parallel reaction monitoring (PRM) or selected reaction monitoring (SRM) (Kang 2012; Gillet et al. 2012). The main steps in LC-MS/MS are:

- (1) Sample preparation and digestion
- (2) Peptide separation using liquid chromatography
- (3) Electrospray ionisation (ESI)
- (4) MS1 (and precursor selection in DDA)
- (5) Precursor ion fragmentation

- (6) MS2
- (7) Peptide identification
- (8) Protein inference

Sample preparation and digestion

Biological samples, particularly plasma or serum, are often immunodepleted of highly abundant proteins to increase the dynamic range of proteins detected (Tu et al. 2010). Off-target depletion of non-abundant proteins and increased variability in quantification are documented unwanted effects of this approach (Tu et al. 2010; Patel et al. 2012; Gunther et al. 2015).

Prefractionation of protein mixtures according to their biochemical properties may be performed prior to digestion, for example using sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) to separate by molecular weight, or isoelectric focusing to separate by isoelectric point (Gundry et al. 2009).

The protein mixture is typically reduced using dithiothreitol (DTT) and alkylated using iodoacetamide (IAA) prior to proteolysis to convert disulphide-bonded cysteine to S-carbamidomethyl cysteine (Gundry et al. 2009). This prevents the reformation of disulphide bonds, allowing protease active sites access to cleavage sites in the peptide backbone.

Hydrolysis of peptide bonds is accomplished using enzymatic cleavage, commonly performed using trypsin. Trypsin is highly efficient and cleaves with high specificity C-terminal to the basic amino acids lysine and arginine (where not followed by proline), thereby producing peptides that are amenable to the chromatographic separation and fragmentation used in typical LC-MS/MS workflows (Tsiatsiani et al. 2015). More recent methods using bead-immobilised heat-stable trypsin to perform proteolysis allow denaturation of proteins using heat (thus avoiding the need for

reduction and alkylation), accelerate proteolysis and may preferentially digest proteins in compartments less well accessed by free trypsin (Shah et al. 2016). Other proteases, with different specificities, such as Lys-C, chymotrypsin or elastase, may be used to alter digestion in specific circumstances (for example, if a particular protein of interest lacks arginine or lysine residues or in the search for post-translational modifications; Tsiatsiani et al. 2015).

Liquid chromatography

Following digestion, peptide mixtures undergo separation using reverse-phase liquid chromatography (RPLC), also referred to as high-performance liquid chromatography (HPLC) or ultra high-performance liquid chromatography (UPLC), according to the particle size of the stationary phase (and consequently operating pressure of the liquid chromatography (LC) system). Peptides are loaded on to the LC column, binding to the hydrophobic stationary phase. They are then eluted from the stationary phase in a gradient of buffer from polar to non-polar solvent, with increasingly hydrophobic peptides eluting later in the gradient. Other chromatographic separations, such as high pH reverse-phase, strong cation exchange or weak anion exchange, may be performed prior to RPLC to provide two-dimensional separation and increase proteomic depth (Keshishian et al. 2015).

Ionisation

Before mass spectrometry is performed, the peptide mixture must be converted into charged ions in the gas phase. In LC-MS/MS this is achieved by coupling RPLC with ESI (Ho et al. 2003). Eluate from LC is passed through a highly-charged capillary, causing the release of positively charged droplets (nebulisation) which, through a combination of repulsion and evaporation of solvent, leads to progressively smaller droplets and finally the release of positively charged ions in the gas phase (Fenn

et al. 1989; Bruins 1998; Ho et al. 2003). The liberated precursor ions are then focused into a beam and guided into the mass analyser by radio frequency (RF) alternating electrical fields produced by quadrupole or octapole ion guides.

Tandem mass spectrometry

In broad terms, tandem mass spectrometry involves the measurement the m/z of all precursor ions entering the mass spectrometer at a given time (MS1), followed by fragmentation of precursor ions. The m/z of these fragment ions is then measured (MS2) and used to infer peptide and protein identification (Kapp et al. 2007).

Mass analysers The measurement of m/z can be performed in several ways. A commonly used instrument in LC-MS/MS is the quadrupole mass filter. Four linear electrodes with alternating RF currents (with opposing polarity in the diagonal rods) generate electrical fields that simultaneously attract and repel ions, causing oscillations according to the m/z of the precursor ion orthogonal to the direction of travel through the analyser. Ions with the selected m/z have a stable trajectory through the analyser, whilst other ions leave the analyser; those with the selected m/z can then be detected or transferred for fragmentation and subsequent MS2 (mass filtering, **Figure 2.1**). Varying the frequency of alternation allows the selection of precursor ions with differing m/z . Stopping potentials can also be applied to the end of a quadrupole to act as an ion trap, though more complex three-dimensional quadrupole ion traps designs are also available (Demartini 2013; March 2009).

Orbitrap analysers record the orbit of ions around and along a spindle-shaped electrode, which is decomposed using a Fourier transform to isolate the frequency of oscillation of different ions, from which the m/z is inferred (Hu et al. 2005; Schaeffer-Reiss 2008). Orbitrap analysers have high resolution and can be used to measure both precursor and fragment ions.

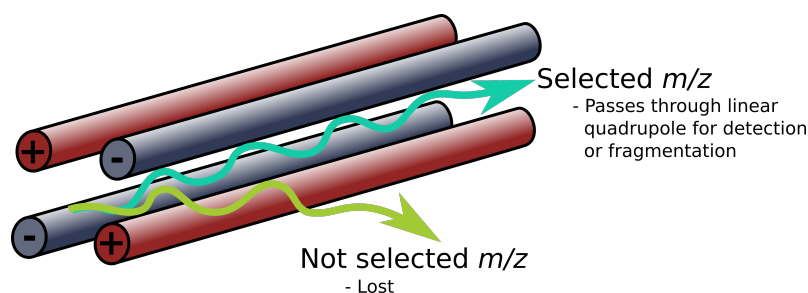


Figure 2.1: Schematic of a linear quadrupole. Diagonal electrodes carry opposing charge. Alternating charge leads to oscillations of ions orthogonal to the direction of travel through the quadrupole. Varying the frequency of oscillation can be used to select ions with specific m/z to pass through the quadrupole for subsequent detection or further analysis such as higher-energy collisional dissociation (HCD). Ions without the selected m/z do not transit the quadrupole.

Modern LC-MS/MS spectrometers often have a combination of mass analysers. The instruments used in this thesis (Q ExactiveTM HF Hybrid Quadrupole-OrbitrapTM and Orbitrap Fusion LumosTM TribridTM mass spectrometers; ThermoFisher Scientific) both have a quadrupole mass filter (used for precursor selection) and orbitrap mass analyser that can be used for both MS1 and MS2 acquisition; the latter instrument also has a dual-pressure linear ion trap, used to measure MS2 spectra in the experiments described in **Chapter 3**. A schematic of the Q ExactiveTM HF Hybrid Quadrupole-OrbitrapTM is shown in **Figure 2.2**.

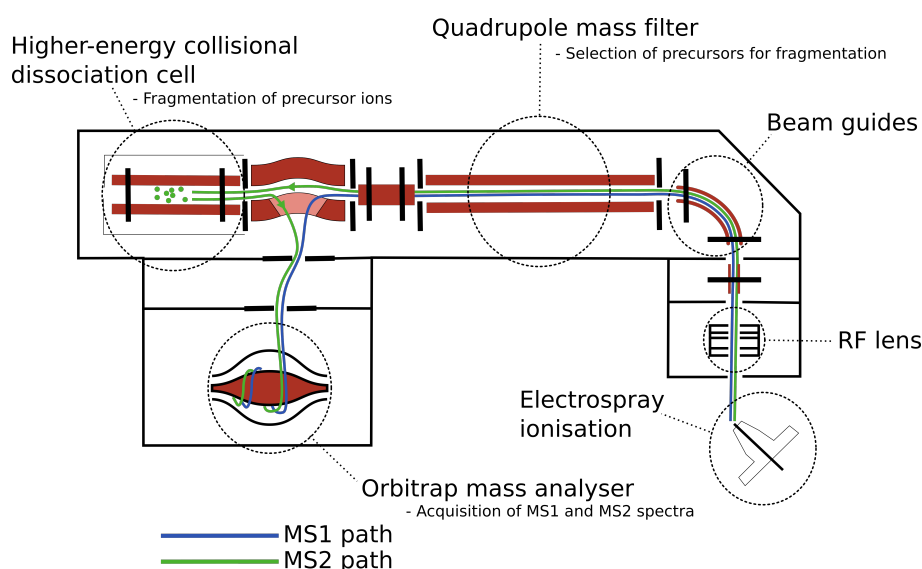


Figure 2.2: Schematic of the Q ExactiveTM HF Hybrid Quadrupole-OrbitrapTM mass spectrometer. Ion paths for MS1 (blue) and MS2 (green) acquisition are indicated. Adapted from ThermoFisher Scientific.

Fragmentation In typical DDA LC-MS/MS workflows, the top N most abundant precursors eluting in a given retention time window are selected by mass filtering for fragmentation and subsequent MS2. In DIA approaches such as sequential window acquisition of all theoretical fragment ion spectra MS (SWATH-MS), *all* precursor ions within a m/z range and retention time window are fragmented followed by acquisition of all MS2 fragments simultaneously which are compared with spectral libraries to aid deconvolution of the constituent precursor ions (Gillet et al. 2012).

Fragmentation of precursor ions takes place in a fragmentation chamber, typically in the form of an ion trap. Collision-induced dissociation (CID) and HCD, the most common means of fragmentation, involve colliding precursor ions with an inert gas, often Argon or Helium, leading to fragmentation of precursor ions (Olsen et al. 2007). Fragmentation mostly occurs along the peptide backbone, usually at the peptide bond site producing complementary b (N -terminus) and y (C -terminus) ions; a precursor ion undergoing fragmentation forms into multiple b - y pairs with a variable number of amino acids along its length (**Figure 2.3** Demartini 2013). Electron transfer and electron capture dissociation methods may be used in certain circumstances, for example to preserve post-translational modifications, and result in the formation of different ion pairs (Demartini 2013). HCD was used in the experiments described in this thesis.

Peptide and protein identification

Peptide identification Differences between the m/z of peaks in the MS2 spectrum correspond to the mass of amino acid residues in the precursor ion (**Figure 2.4**). Multiple database search algorithms are available to calculate the predicted amino acid sequence from this data and map this to peptides and subsequently proteins, including Mascot (Matrix Science; used herein), Comet (University of Washington), PEAKS DB (Bioinformatics Solutions Inc) X!Tandem (The Global Proteome Machine Organization), and others (reviewed in Cottrell 2011). Broadly,

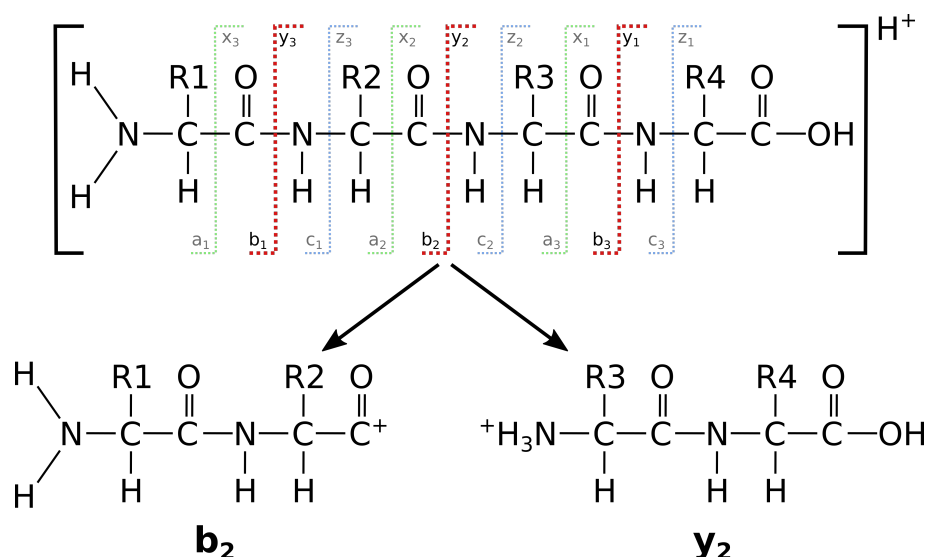


Figure 2.3: Fragmentation of precursor ions. Fragmentation in HCD and CID typically results in b-y ion pairs, whose m/z can be used to infer the peptide sequence of the precursor ion. Other fragmentation methods such as electron capture and electron transfer dissociation lead to the formation of different ion pairs.

identification of the precursor ion is performed by comparing the amino acid sequence obtained from MS2 to peptides predicted by *in silico* digestion of the appropriate protein database (such as the UniProt human reference proteome; Cottrell 2011). Database searching must take into account mass accuracy of the instrumentation, fixed modifications (such as carbamidomethylation of cysteine, if reduction and alkylation have been used; or isobaric labelling) and variable modifications (such as post translational modifications) and the specificity of the enzyme used for digestion (this may also allow for missed cleavages; Cottrell 2011).

Database searching algorithms use scoring systems to denote the confidence of a peptide spectral match (PSM), assigning the highest scoring peptide to a particular spectrum, though scoring is not standardised between algorithms (Sadygov et al. 2004). Estimating the FDR of peptide hits is frequently used to estimate the proportion of false positive peptides in a dataset (Elias et al. 2007). This is usually achieved using a target-decoy methodology, by searching spectra against a database of target proteins concatenated with a decoy database made by reversing the sequence of target proteins (Elias et al. 2007). The FDR of peptide hits can then be estimated, based on the number of false positive hits (i.e. hits within the decoy

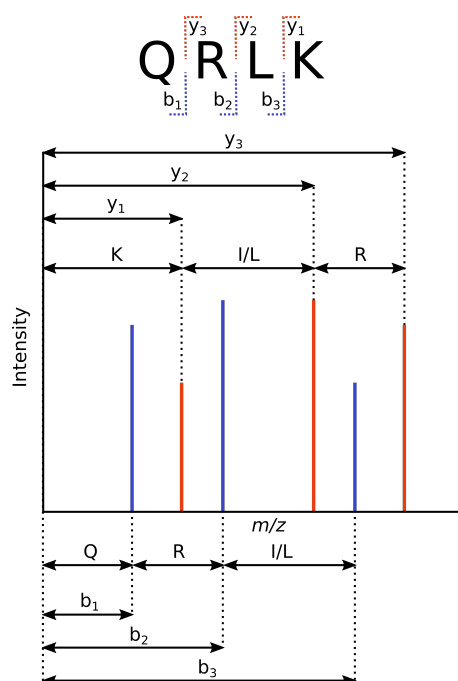


Figure 2.4: A simplified MS2 spectrum for the peptide QRLK. Differences in the m/z of fragment ions can be used to infer the amino acid sequence of a precursor peptide (note that leucine and isoleucine have the same molecular mass therefore cannot be distinguished).

database), and the peptide score threshold adjusted to produce a predefined peptide FDR (typically 1%).

Protein inference Most PSMs will not be unique to a protein due to redundancy and protein sequence homology (Nesvizhskii et al. 2005). Protein inference describes the process by which database search algorithms attempt to ascertain which proteins were present in a sample, given the set of PSMs identified. Where a peptide is unique to one protein in the target database, it is simple to say that that protein is present; peptides that overlap between homologous proteins are more problematic. Different search algorithms handle conflicting peptides differently, but generally apply Occam’s razor to collapse proteins indistinguishable by their PSMs to a single protein group (Nesvizhskii et al. 2005).

Quantification

Label-free quantification Estimating the relative abundance of proteins can be achieved without additional labelling of samples using MS1 or MS2 data. MS1 quantification methods use the intensity of precursor ions, which has been shown to vary linearly with protein abundance (Megger et al. 2013). MS2 methods use spectral counting, based on the stochastic nature of ion selection for fragmentation, or the intensity of MS2 spectra. MS1 intensity-based methods generally outperform spectral counts in the accuracy of quantification, though performance using MS2 intensities is comparable (Grossmann et al. 2010; Krey et al. 2014; Bubis et al. 2017). Label-free methods require LC-MS/MS be performed separately on each sample; MS2 quantification methods necessitate identification of a protein or peptide in every sample in order to compare its abundance, whereas MS1 intensity methods allow the matching of precursor ion peaks between LC-MS/MS runs, though shifts in retention time may influence the accuracy of matching (Cox et al. 2014).

Labelling methods for biomarker discovery: isobaric tags Labelling of peptides for mass spectrometry can be achieved using isobaric tags reactive for the amine group. These tags comprise a reporter group, with differing molecular mass, and a balance group, the molecular mass of which is varied to make the overall mass of the tags identical (Rauniyar et al. 2014). Samples are labelled following digestion, after which they are combined for prefractionation and LC-MS/MS. Reporter ions are liberated during fragmentation, giving peaks on the MS2 spectrum with intensity that corresponds to relative abundance. Isobaric tags allow the simultaneous measurement of several (up to 11) samples, thereby reducing LC-MS/MS time, are more accurate in quantification than label-free methods, but generate fewer identifications than label-free methods (Li et al. 2012).

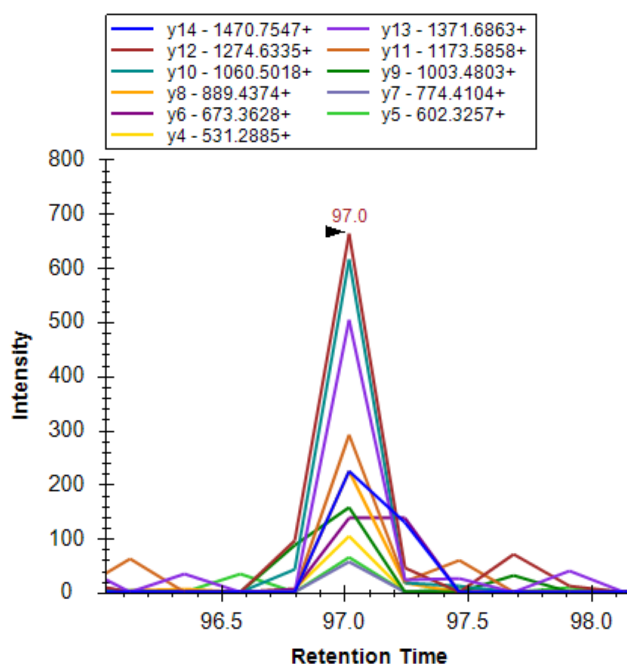


Figure 2.5: Parallel reaction monitoring of y_4 - y_{14} fragments of a 20-amino acid precursor ion of Glutamate receptor 4 (GRIA4) in CSF.

Targeted proteomics

Targeted proteomic analysis methods, specifically SRM, multiple reaction monitoring (MRM) and PRM, monitor the fragmentation spectra of defined precursor ions (precursor-fragment ion pairs known as transitions) over time to measure peptide abundance (Lange et al. 2008; Kuhn et al. 2004; Anderson et al. 2006). Precursor ions are selected according to retention time and precursor m/z , fragmented and the intensity of fragment ions is measured over time, with the integrated intensity or maximal intensity used to infer abundance (**Figure 2.5**).

Spiking samples with stable heavy isotope-labelled synthetic peptides enables absolute quantification of peptides, though this may not accurately represent protein abundance, for example due to incomplete digestion (Gerber et al. 2003). Production of heavy-labelled proteins (PSAQ) or artificial protein composed of concatenated target peptides (QconCAT) is a further step which circumvents these issues (Pratt et al. 2006). Such heavy proteins are spiked into samples prior to digestion,

negating incomplete digestion as a cause of inaccurate quantification in targeted proteomic experiments.

2.1.3 Data analysis

Missing data imputation

Missing data is problematic in DDA label-free proteomic analysis; although more problematic when using MS2-level label-free quantification (where peak-matching is not used), missing data also occurs with peak-matched MS1 intensity-based quantification. The underlying causes of missing data in label-free proteomic experiments are complex, and can occur at any stage in the experiment from sample degradation, missed proteolytic cleavage, ion suppression, aberrant peak matching and peptide misidentification (Webb-Robertson et al. 2015). As such, missing proteomic data is considered missing not at random (MNAR) if due to low abundance, or missing at random (MAR) if missing for other reasons.

Numerous methods have been developed to impute missing data using the distribution of present data, predominantly applied to overcome data either MNAR or MAR, though the underlying cause of missingness is rarely known for individual proteins (Lazar et al. 2016). Single-value imputation methods such as deterministic or probabilistic minimum imputation (i.e. imputation close to the minimum measured value) are designed to deal with for MNAR values, whilst local or global similarity approaches (such as maximum likelihood estimator (MLE), k -nearest neighbours (k NN) or singular value decomposition (SVD) imputation) are more appropriate for MAR values and generally perform better except where there is a high proportion of MNAR values (Lazar et al. 2016).

During the work in this thesis, imputation by k NN was selected as it performs well imputing at both the peptide and protein level with little missing data and was not

disruptive to the distribution of protein intensities in the datasets presented (Lazar et al. 2016).

Normalisation

Due to differences in the protein content loaded on to LC-MS/MS, data normalisation is necessary to avoid inflated variance of protein measurements. Multiple normalisation strategies have been described (often originally developed for microarray experiments). Frequently-employed approaches include log normalisation with median centering, with or without scaling by standard deviation (SD) or median absolute deviation (MAD), variance stabilising transformation (which attempts to make sample variances independent of mean intensity), quantile normalisation (which replaces values for each quantile with the mean intensity of that quantile across all data) and linear regression (Bolstad et al. 2003; Callister et al. 2006; Motakis et al. 2006; Valikangas et al. 2018).

Statistical analysis of proteomic data

Label-free quantitative data can be analysed using univariate or multivariate approaches (Marengo et al. 2014). The univariate approach is suitable for the generation of candidate biomarkers can be used to generate lists of differentially regulated proteins that can then be used in analysis of perturbations in the protein network, such as protein-protein interaction (PPI) network analysis, GO enrichment analysis or pathway analysis (Schmidt et al. 2014). Univariate analysis is accomplished by testing the abundance of individual protein groups using parametric tests such as two-sample t -tests (or f -tests), where assumptions of normality of data are not violated. Where distributions are non-normal, alternatives are to use non-parametric tests (such as the Mann-Whitney U test, although this is problematic when n is small) or to test against an empirical distribution using permutation testing. Alter-

natives to conventional parametric tests have been developed for use in genomics and may also be applied to proteomics, such as reproducibility-optimised test statistic (ROTS) and empirical Bayes-adjusted t -tests (which adjust each protein's variance towards the pooled variance) as implemented in the R package LIMMA (Smyth 2004; Elo et al. 2008; Kammers et al. 2015).

Importantly, due to the number of statistical tests performed in such univariate proteomic analysis, adjustment is necessary to account for the uniform null distribution of p -values. This is achieved by controlling either the family-wise error rate (FWER) or FDR to a pre-defined level. FWER-controlling procedures, such as the Holm-Bonferroni correction, control the likelihood of *any* false positive results and are thus highly conservative (Holm 1979). More recent developments such as the Benjamini-Hochberg step-up procedure and Storey q -value, provide FDR control (estimating the proportion of false positives at a given significance level) and are therefore less conservative and more widely used in proteomic and genomic sciences (Benjamini 1995; Storey 2003).

As with other high-dimensional datasets, proteomic analysis lends itself well multivariate analysis techniques. Unsupervised methods such as principal component analysis (PCA) and hierarchical clustering are frequently employed to explore patterns within proteomic datasets and observe clustering of samples (Marengo et al. 2014). A broad range of more complex machine learning techniques for supervised classification (such as partial least squares discriminant analysis (PLS-DA), support vector machine and random forest) and regression models may also be used to identify features with diagnostic or prognostic potential.

2.2 Size exclusion chromatography

The separation of molecules by size (including proteins) within a mixture using porous materials (known as size exclusion or gel chromatography) is long-established. The experiments reported in **Chapters 5** and **6** use size exclusion chromatography as the primary means by which to separate EVs from other fluid constituents. The principles of size exclusion chromatography are described below.

2.2.1 Basic principles of size exclusion chromatography

In size exclusion chromatography, columns are packed with matrix (also known as resin) comprising porous spherical particles. Various resin materials are available, including dextrans (such as Sephadex), agarose and polyacrylamide (Hong et al. 2012). The key properties that make a material suitable for use as size exclusion resin are inertness, low adsorption and an appropriate range of pore sizes for the chosen application. For the EV extraction described in **Chapters 5** and **6**, Sepharose 4 Fast Flow (a cross-linked agarose matrix; GE Healthcare Life Sciences, UK) was used due to its tolerance for high flow rates at moderate pressures and high exclusion limit (M_r 3×10^7), allowing separation of EVs from large proteins.

This "bed" is equilibrated with buffer, which fills the pores of the matrix (constituting the stationary phase) and surrounds the particles (constituting the mobile phase). Sample is applied to the column and eluted in buffer; the eluate volume at which sample constituents elute from the column is dependent upon size (or more strictly Stokes radius). Small molecules diffuse extensively within the stationary phase and therefore take longer to elute from the column, whilst larger molecules diffuse less within the stationary phase and elute earlier. Sample constituents that are larger than the maximum pore size will not diffuse within the stationary phase at all and therefore elute first, in the void volume (corresponding to the volume of

the mobile phase; GE Healthcare Life Sciences, 2014; Hong et al. 2012).

Various parameters can be tuned to improve the resolution of size-exclusion chromatography, including flow rate (slower flow improving resolution), sample load (sample volume $<4\%$ column volume enabling high-resolution fractionation) and column dimensions (longer columns tend to improve resolution, but smaller column volumes permit smaller eluate volumes; GE Healthcare Life Sciences, 2014; Hong et al. 2012).

For the purposes of CSF EV extraction, bulk fractionation would have been adequate (simply separating EVs from smaller sample constituents, i.e. proteins). However, optimisation to produce small EV-containing eluate volumes (and therefore restrict the amount of concentration required prior to digestion) led to the combination of small column volume (26mLs) and relatively small sample volume (800 μ L, 3.07% column volume) that would be consistent with high-resolution fractionation.

2.3 Participants and sampling

2.3.1 Ethical Approval and Subject Recruitment

Ethical approval for work in this thesis was obtained from South Central Oxford Ethics Committee B (08/H0605/85), South Central Oxford Ethics Committee C (09/H0606/5+5), NRES Central Committee South Central – Berkshire (14/SC/0083 and 10/H0505/71). All participants provided written consent (or gave permission for a carer to sign on their behalf). Patients with ALS, PLS and mimic disorders were recruited from the Oxford MND Centre, Oxford, UK and patients with Parkinson’s disease were recruited through the Oxford Parkinson’s Disease Centre, Oxford UK. Healthy control subjects were typically spouses and friends of those with ALS, PLS and mimic disorders.

CSF samples used in the development of methods of CSF proteomic analysis and CSF EV extraction in **Chapter 5** were obtained from patients undergoing lumbar puncture for the investigation of neurological symptoms in the Neurosciences Ward, Oxford University Hospitals NHS Foundation Trust (Ethics Committee South Central – Oxford C, reference 09/H0606/5+5 and NRES Committee South Central – Hampshire B, reference 14/SC/1248).

Exclusion criteria included age below 18 years and or contraindication to lumbar puncture including inherited bleeding disorders, chronic liver disease and use of anticoagulant or antiplatelet drugs, and systemic illness at the time of CSF sampling. Control subjects were excluded if there was any family history of ALS or PLS. Patients with ALS, PLS and mimic disorders were diagnosed according to standard criteria by two neurologists experienced in the diagnosis of motor disorders in the Oxford clinic (Brooks et al. 2000b; Gordon et al. 2006).

CSF was collected at baseline and, in ALS and PLS patients, every 3-6 months when available up to a maximum of five occasions. Patients with ALS and PLS were followed until death or censorship.

2.3.2 Biofluid Sampling and Processing

Biofluid samples were processed and stored in accordance with consensus guidelines for biomarker development (Teunissen et al. 2011).

CSF samples were obtained by lumbar puncture from the L3-L4 interspace with the subject positioned seated. The skin was cleaned with chlorhexidine wash prior to the procedure. Up to 2 mLs local anaesthetic (lignocaine 1%) was infiltrated subcutaneously. CSF was obtained using a 22 gauge Sprotte[®] or Quincke needle. 10-14 mLs CSF was extracted into 25 mL capacity polypropylene tubes. Within one hour of sampling, CSF samples underwent centrifugation at 3000*g* for 10 mins

at 4°C, the supernatant collected and frozen at -80°C until use.

Twenty-five mLs of venous blood was obtained from each participant using the BD Vacutainer® Safety-Lok™ blood collection set into 5.0 mL serum separator tubes. Blood samples underwent centrifugation at 3000*g* for 10 mins at 4°C, serum was collected and frozen at -80°C until use.

2.3.3 Clinical parameters

Symptom onset was defined as first weakness (for ALS or PLS) or relevant neurological symptom reported by patients. Degree of disability for patients was obtained using the ALSFRS-R (48-0) at each research visit (Moore et al. 2004). Cognitive impairment was sought using the Edinburgh cognitive and behavioural ALS screen (ECAS), a cognitive score specifically validated for identification of executive dysfunction in ALS that can be administered independent of difficulty with speech or writing (Abrahams et al. 2014). ECAS score was obtained for each participant at their first research visit.

PR was calculated per-visit using the ALSFRS-R by $[48 - \text{ALSFRS-R}] / [\text{months from symptom onset}]$ (Labra et al. 2016). Patients were divided into fast- and slow-progressing using a PR above or below the median value. A clinical upper motor neuron burden score (0-15) was calculated by the presence of hyperreflexia in each of 15 reflexes (in the face, glabellar and facial reflexes and jaw jerk; in the upper limb, triceps and biceps reflexes and finger jerks; in the lower limb knee and ankle reflexes and plantar responses), obtained at their first research visit (Turner et al. 2004).

2.4 Statistical analysis

Statistical analysis was performed and plots produced using R v3.3.3 (Team 2015). Normality of data was assessed using visual inspection of a normal quantile-quantile plot. Univariate comparisons of relative and absolute abundance were performed using Welch's t -test for normal (or log-normal) data unless otherwise specified. For non-normally distributed data, the Mann-Whitney U test was used. Comparisons between multiple groups were achieved using pairwise t -tests or Mann-Whitney U tests with FDR correction. FDR correction was performed using the Benjamini-Hochberg step-up procedure (BH) unless otherwise specified. FDR-adjusted or raw $p < 0.05$ were used as appropriate to denote statistical significance.

For shotgun proteomic experiments, statistical inference was performed on protein data. Protein abundance values were centered to the median abundance of the 90% of proteins with the lowest variance across all runs and scaled by median absolute deviation (Keilhauer et al. 2015). Missing data were imputed using 3 nearest neighbours imputation, unless otherwise specified.

Receiver operator characteristic (ROC) analysis was performed to assess classifier performance and ROC curves were compared using the DeLong method (using paired analysis to account for the use of multiple markers in the same samples; DeLong et al. 1988). Sensitivity and specificity are given for cut-offs at the Youden index (Youden 1950).

3. The Cerebrospinal Fluid Proteome in ALS

3.1 Hypothesis and Aims of This Chapter

3.1.1 Hypothesis

Changes in the ALS CSF proteome reflect underlying disease processes and can provide objective measures of disease activity.

3.1.2 Aims

1. Perform proteomic analysis of CSF of patients with ALS and relevant control groups using LC-MS/MS
2. Analyse the CSF proteome in cross-sectional samples to identify candidate markers
3. Examine the performance of identified candidate markers in terms of:
 - i. Within-sample classifier performance
 - ii. Correlation with established clinical parameters
 - iii. Longitudinal stability
 - iv. Association with survival
4. Validate candidate proteins identified in previous ALS CSF proteomic analyses

3.2 Introduction

Although previous proteomic studies of the CSF of patients with ALS, detailed in **Section 1.3.1**, have yielded some overlap in differentially abundant proteins, no *consistent* aberrations have been observed across experiments (Barschke et al. 2017). Although this might reflect differences in the patient and control cohorts used, this variability is in part attributable to methodological differences adopted at all stages. Particular variation has occurred in sample preprocessing, with studies variously using abundant protein depletion and prefractionation by SDS-PAGE or LC-based fractionation. The choice of mass spectrometry method adopted has also varied greatly (with early studies using SELDI-TOF of intact proteins and more recent studies using LC-MS/MS).

Importantly, most studies have used sample pooling. Pooling may be advantageous in some circumstances, since it can overcome issues such as a lack of individual sample volume, avoid excessive MS instrument time and reduce biological variability (thereby increasing statistical power for a given number of replicates; Diz et al. 2009). However, there are multiple issues with sample pooling that severely limit its usefulness. Although protein abundance in a pooled sample ought theoretically to represent the mean abundance, sample pools are susceptible to biases due to pipetting error for which adjustment cannot be made *post hoc* (Oberg et al. 2009). In addition, the averaging that occurs during sample pooling is not on the log scale, typically employed for the analysis of protein abundance and is therefore not equivalent to averaging log abundance from multiple individual samples (Oberg et al. 2009).

Sample pooling also leads inevitably to loss of phenotypic data, therefore limiting clinical-biochemical associations that can be sought. Grouping of samples according to clinical data partially circumvents this, though the multidimensional heterogeneity inherent to ALS renders any sample pooling unsatisfactory.

For these reasons, an approach that minimises preanalytical sample processing and enables individual sample proteomics was selected for this study, utilising a plate-based bead-bound heat-stable trypsin to study CSF from patients with ALS and relevant control groups.

3.3 Materials and methods

3.3.1 Participants and sampling

The study included 43 patients with ALS, 6 patients with PLS, 20 patients with Parkinson's disease, 12 patients with mimic conditions (including multifocal motor neuropathy with conduction block, Kennedy's disease, cervical spondylotic myelopathy and motor presentation of multiple sclerosis) and 20 healthy control subjects. Subject recruitment, CSF sampling, sample processing and clinical characterisation was completed as described in **Section 2.3**. Demographic and clinical characteristics are displayed in **Table 3.1**.

3.3.2 Sample preparation for proteomic analysis

Samples of CSF were thawed on ice and digested using heat stable immobilized trypsin (SMART DigestTM, ThermoFisher Scientific, UK). Fifty μL of CSF was mixed with 150 μL SMART DigestTM buffer and added to SMART DigestTM plates. Samples were incubated at 70°C with shaking at 1400 rpm for 60 minutes. Digested samples were desalted using SOLA μ TM solid-phase extraction (SPE) plates (ThermoFisher Scientific, UK) and dried by vacuum centrifugation. Samples were resuspended in 20 μL buffer A (2% acetonitrile, 0.1% formic acid in water) and kept at -20°C until analysis. Peptide concentrations were assayed using a PierceTM quantitative colorimetric peptide assay (ThermoFisher Scientific, UK) according to

	ALS	PLS	Healthy control	Parkinson's disease	Mimic conditions	<i>p</i>
n, first visit	43	6	20	20	12	-
n, second visit	21	5	-	-	-	-
n, third visit	13	3	-	-	-	-
n, fourth visit	10	2	-	-	-	-
n, fifth visit	2	-	-	-	-	-
% male	72.1	16.7	55.0	50.0	91.7	0.010*
Age at sampling, mean $\pm$ SD (years)	61.9 $\pm$ 10.3	68.0 $\pm$ 6.9	58.5 $\pm$ 8.6	62.6 $\pm$ 4.1	57.7 $\pm$ 15.5	0.180 ⁺
Age at onset, mean $\pm$ SD (years)	59.3 $\pm$ 10.9	40.7 $\pm$ 34.4	-	60.8 $\pm$ 4.0	-	0.766 ⁺
% bulbar onset	18.60	16.67	-	-	-	0.140 [†]
Disease progression rate median [IQR]	0.50 [0.28-0.95]	0.12 [0.09-0.15]	-	-	-	0.001 [‡]
Duration of follow-up, mean $\pm$ SD (months)	34 $\pm$ 26	80 $\pm$ 15	-	-	-	< 0.001 [‡]
Deaths, n (%)	35 (81)	0	-	-	-	< 0.001 [†]

Table 3.1: Demographic data. * χ^2 test with *post hoc* pairwise Fisher exact tests with FDR-adjusted $p < 0.05$ for mimic vs PLS only. ⁺ Kruskal-Wallis H test. [†] Fisher exact test. [‡] Mann-Whitney U test. ALS – amyotrophic lateral sclerosis; PLS – primary lateral sclerosis.

the manufacturer's instructions. A pooled sample was produced by combining equal quantities of digested peptide from each individual sample and injected after every ninth sample for use in quality control analysis.

3.3.3 Proteomic analysis

Peptides were analysed by nano ultra-high performance liquid chromatography tandem mass spectrometry (nUHPLC LC-MS/MS) using a Dionex UltiMate 3000 ultra high performance liquid chromatography (UHPLC) (ThermoFisher Scientific, Germany) coupled to a Q ExactiveTM HF Hybrid Quadrupole-OrbitrapTM mass spectrometer (ThermoFisher Scientific, Germany). Five hundred nL of peptides from each sample was injected and analysed using a 60-minute linear gradient at a 250 nL/minute flow rate. The gradient used to elute the peptides started at 3 min-

utes with 2 % buffer B (0.1% trifluoroacetic acid (TFA) and 5% dimethyl sulfoxide (DMSO) in acetonitrile) increasing to 5% by 6 minutes followed by an increase up to 35% by 63 minutes. The data were acquired with a resolution of 60,000 full-width at half maximum (FWHM) ion intensity with a m/z ratio of 400 and lockmass enabled at 445.120025 Thomson (Th). The 12 most abundant precursors in each MS1 scan were selected for fragmentation by HCD at a normalized collision energy of 28 followed by exclusion for 27 seconds.

Raw MS data were analysed using Progenesis QI for Proteomics software v3.0 (Non-linear Dynamics). MS/MS spectra were searched against the UniProt Homo Sapiens Reference proteome (retrieved 01/06/2017) using Mascot v2.5.1 (Matrix Science) allowing for a precursor mass tolerance of 10 ppm and a fragment ion tolerance of 0.05 Da. Deamidation on asparagine and glutamine and oxidation on methionine were included as variable modifications. The peptide FDR was set at 1% and all peptides with an ion score higher than 20 were imported into Progenesis QIP. Proteins that were defined with at least one unique peptide were included in the protein data set for further analysis. Protein abundance values were centered to the median abundance of the 90% of proteins with the lowest variance across all runs and scaled by MAD (Keilhauer et al. 2015). Missing data were imputed using three nearest neighbours imputation.

3.3.4 ELISA

CSF pNFH was measured using a commercially-available ELISA for pNFH (Euroimmun AG, Germany). CSF samples were assayed neat. 100 μ L enzyme conjugate and 25 μ L of standards and samples were added to the plate in duplicate and incubated for 2 hours at room temperature with 400rpm shaking. The plate was washed 3 times with 300 μ L wash buffer. 100 μ L substrate reagent was added and incubated at room temperature for 15 minutes with 400rpm shaking protected from

light. 100 μ L stop solution was added and plates were read immediately at dual wavelengths (450-620nm). Plates were read using a FLUOstar Omega plate reader (BMG LABTECH, UK). Standard curves were fitted with 4-parameter logistic regression using MARS data analysis software (BMG LABTECH, UK). CSF samples were thawed on ice prior to measurement.

The median intra-assay coefficient of variation (CV) was 2.54% (interquartile range (IQR) 1.38-5.52%) and interassay CV was 7.20% (IQR 3.54-12.86%). The lower limit of quantitation (LLOQ) was 0.32 ng/mL.

3.3.5 Statistical analysis

Statistical analysis was performed as described in **Section 2.4**. Cross-sectional analysis was performed using data and samples obtained from an individual's first visit. Due to the non-normal distribution of normalised log protein abundances, a Mann-Whitney U test was used to test for changes in abundance between ALS samples and those from each control group. Correlation with clinical parameters was calculated using Pearson's correlation coefficient and associations between predictor and outcome variables tested using Student's t -test. Correlation with phosphorylated neurofilament heavy chain (pNFH) levels was performed using data across all visits. Cox regression analysis of survival was performed incorporating known survival covariates and univariate analysis using the log-rank test, again using data from first visit samples. Longitudinal data analysis was performed using a random intercept, random slope linear mixed model. Comparison of candidate markers identified in previous proteomic studies was performed between ALS and other groups separately using Welch's t -test with Benjamini-Hochberg adjustment for multiple comparisons.

3.4 Results

3.4.1 Proteomic analysis performance

Seven hundred seventy-three protein groups were identified and quantified across conditions with at least one unique peptide. Median CV of pooled samples was 12.48% (IQR 7.93–23.4) across all protein groups quantified. Median combined CV (biological and analytical) across all protein groups in first visit samples from all subjects was 20.11% (IQR 14.68–28.82; difference in variation between pools and individual first visit samples, $p < 0.001$), with 74.11% of proteins having CV $< 25\%$, giving a statistical power of at least 0.8 for a 1.5-fold difference in these proteins at FDR-adjusted $p < 0.05$ for all comparisons (Levin 2011). Quantification of pooled samples was highly correlated, with median between-pool correlation Pearson’s $r = 0.94$ $p < 0.001$ for all between-pool correlations; IQR, 0.90–0.97). In relation to the three candidate markers identified, combined variation in first samples across all subjects was 45.86% for CHIT1 (combined variation in ALS samples 45.73%, healthy controls 24.39%, PD 25.22%, PLS 30.21%, and mimic 27.19%), 16.21% for CHI3L1 (combined variation in ALS samples 15.18%, healthy controls 12.67%, PD 12.75%, PLS 3.55%, and mimic 14.32%), and 36.8% for CHI3L2 (combined variation in ALS samples 38.49%, healthy controls 21.73%, PD 28.22%, PLS 35.09%, and mimic 33.01%). The within-individual variation across multiple visits for individuals undergoing longitudinal sampling was median 12.52% (IQR, 7.06–21.87) for CHIT1, 4.79% (2.70–6.92) for CHI3L1, and 14.27% (8.97–22.62) for CHI3L2.

To examine changes in the accuracy of quantification across the all samples, normalized log abundances of all protein groups in the first sample were subtracted from the subsequent samples and a linear model created to examine changes in the median deviation of relative abundance of all proteins from the first sample. This did not change significantly over the course of the experiment (gradient = 0.00; log

abundance units per sample, $p = 0.062$).

3.4.2 Cross-sectional candidate biomarker identification

CHIT1, CHI3L1 and CHI3L2 are elevated in ALS

Nineteen protein groups were differentially abundant with FDR-adjusted $p < 0.05$, 3 between ALS and healthy controls, 16 between ALS and Parkinson’s disease and 1 between ALS and mimic disorders. No proteins were differentially abundant with FDR-adjusted $p < 0.05$ between ALS and PLS (**Figure 3.1**).

The proteins with FDR-adjusted $p < 0.05$ between ALS and healthy controls comprised three chitinases produced by cells of the immune system, particularly macrophages: CHIT1, CHI3L1, and CHI3L2 (Di Rosa et al. 2016). All three proteins were elevated in ALS compared to other groups except PLS. CSF levels of CHI3L1 did not differ significantly between ALS and PLS ($\log_2$ fold change 0.04; unadjusted $p = 0.666$), though elevation of CHIT1 and CHI3L2 were observed in ALS compared to PLS (**Figure 3.2**; fold change, unadjusted p -value, and FDR-adjusted p -value for chitinase proteins are given in **Table 3.2**).

Protein	ALS vs HC			ALS vs PD			ALS vs PLS			ALS vs mimic		
	FC	p	FDR	FC	p	FDR	FC	p	FDR	FC	p	FDR
Chitotriosidase 1	0.87	<0.001	<0.001	0.85	<0.001	<0.001	0.62	0.014	0.755	0.64	<0.001	0.197
Chitinase-3-like protein 2	0.43	<0.001	0.005	0.39	0.002	0.025	0.59	0.002	0.471	0.62	<0.001	0.081
Chitinase-3-like protein 1	0.23	<0.001	0.013	0.23	<0.001	0.011	0.04	0.666	0.915	0.32	<0.001	0.004

Table 3.2: Relative abundance of chitinase proteins in CSF from patients with ALS compared with healthy controls and those with Parkinson’s disease, PLS and mimic conditions. ALS – amyotrophic lateral sclerosis; PLS – primary lateral sclerosis; FC – $\log_2$ fold change; p – unadjusted p -value; FDR – false discovery rate.

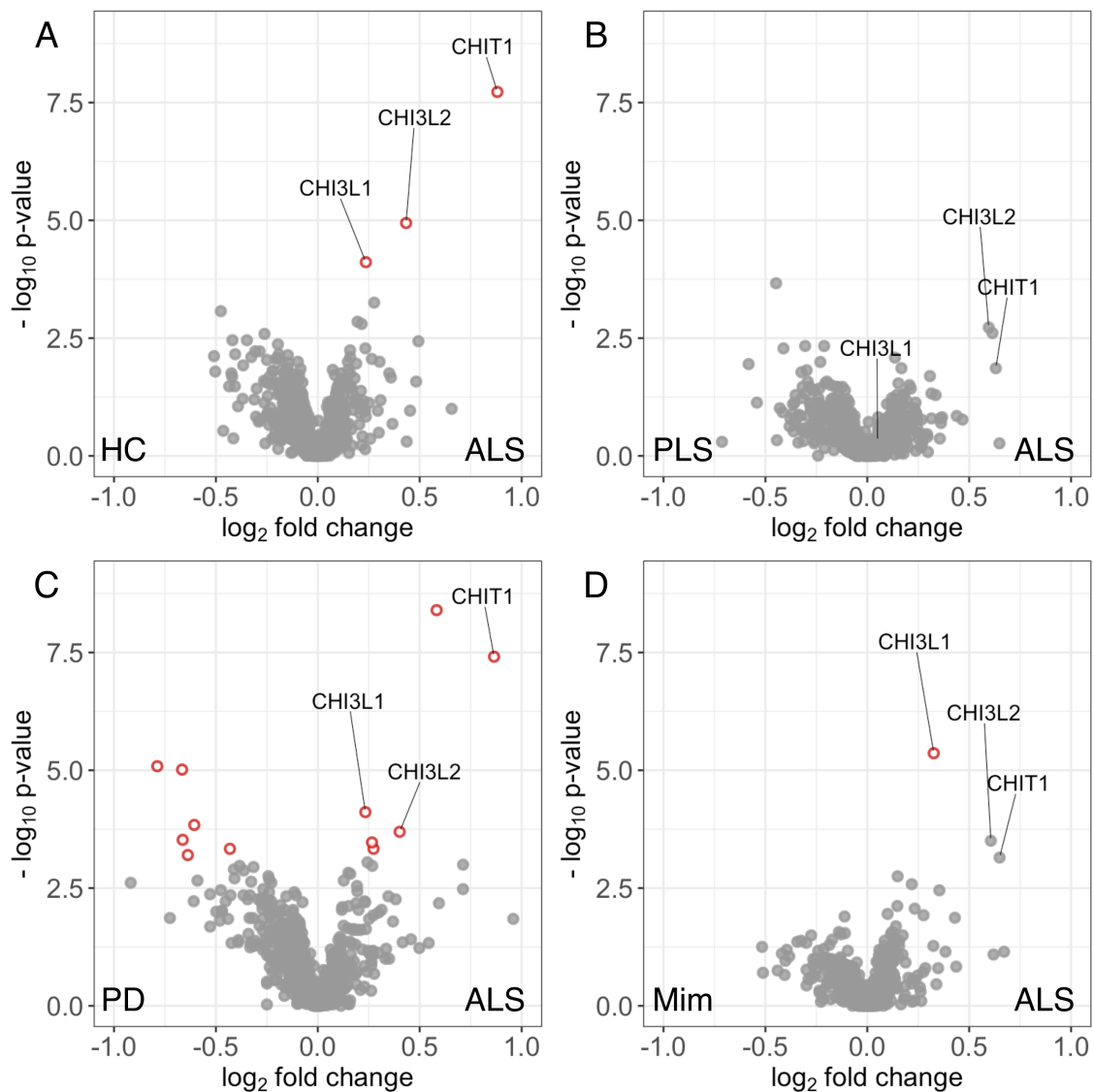


Figure 3.1: Volcano plot showing $\log_2$ fold change vs unadjusted p -values for 773 proteins quantified using LC-MS/MS with label free quantification. Positive $\log_2$ fold change indicates proteins upregulated in ALS. Unfilled red points represent FDR-adjusted p -values < 0.05 . Consistently differentially abundant proteins CHIT1, CHI3L1 and CHI3L2 are labelled. **A** ALS vs healthy control subjects **B** ALS vs PLS **C** ALS vs Parkinson's disease **D** ALS vs mimic disorders. ALS – amyotrophic lateral sclerosis; HC – healthy control; PLS – primary lateral sclerosis; PD – Parkinson's disease; CHIT1 – chitotriosidase 1; CHI3L1 – chitinase-3-like protein 1; CHI3L2 – chitinase-3-like protein 2.

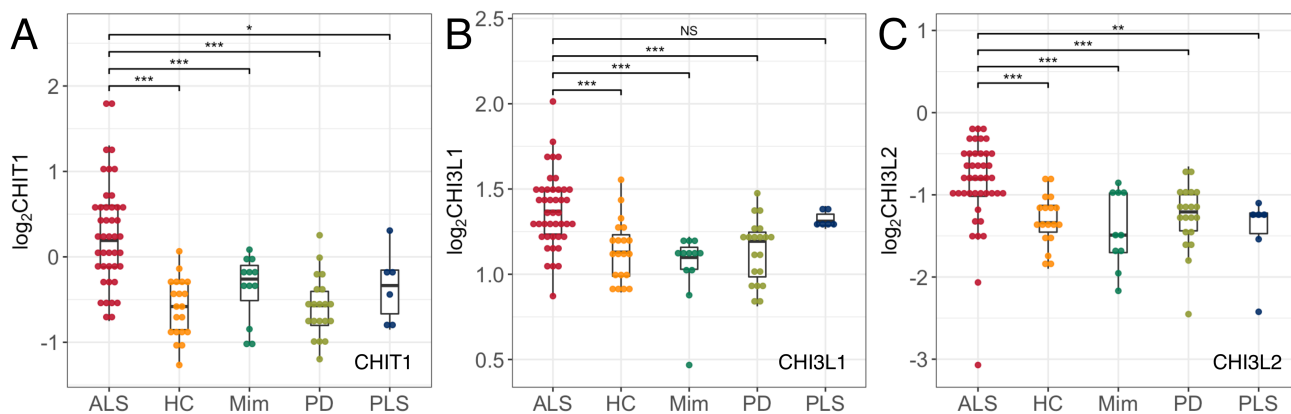


Figure 3.2: Cross-sectional analysis. Normalised $\log_2$ relative abundance of **A** CHIT1, **B** CHI3L1 and **C** CHI3L2 between conditions (first visit samples only). Significance bars indicate unadjusted p -values for first visit samples. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS $p > 0.05$. ALS – amyotrophic lateral sclerosis; HC – healthy control; Mim – mimic disorders; PLS – primary lateral sclerosis; PD – Parkinson’s disease; CHIT1 – chitotriosidase; CHI3L1 – chitinase-3-like protein 1; CHI3L2 – chitinase-3-like protein 2.

3.4.3 Classifier performance of identified candidate markers

Multivariate ROC analysis was performed to assess the utility of the three markers in differentiating ALS from other sample groups (using first visit samples only). **Table 3.3** details the ROC AUC for distinguishing ALS from the other conditions for each marker as well as a combination of all three markers, calculated by summation of centered and squared log abundances. ROC curves for this data are illustrated in **Figure 3.3**.

Differentiating ALS from mimic conditions and PLS are the two most clinically relevant classification problems. Performance distinguishing ALS and mimic conditions was good for all markers (CHIT1 AUC 0.81, 95% confidence interval (CI) 0.69-0.93; CHI3L1 AUC 0.90, CI 0.82-0.98; CHI3L2 AUC 0.81, CI 0.68-0.95), though a combination of all three markers (by summation of centered and scaled log abundance values) provided the highest AUC, improving significantly upon the performance of CHIT1 and CHI3L2, but not CHI3L1 alone (combined chitinases AUC 0.95, CI 0.88-1.00; comparison with CHIT1 alone $p = 0.049$, CHI3L1 $p = 0.384$, CHI3L2 p

	ALS vs healthy control				ALS vs Parkinson's disease			
	AUC	CI	Spec.	Sens.	AUC	CI	Spec.	Sens.
Chitotriosidase 1	0.90	0.83-0.98	0.95	0.77	0.90	0.82-0.98	0.90	0.78
Chitinase-3-like protein 1	0.80	0.68-0.92	0.80	0.74	0.80	0.69-0.91	0.85	0.67
Chitinase-3-like protein 2	0.83	0.73-0.94	0.83	0.81	0.79	0.67-0.90	0.70	0.81
Combined chitinases	0.93	0.87-0.99	0.94	0.76	0.92	0.85-0.99	0.70	0.98
pNFH	0.98	0.95-1.00	1.00	0.93	-	-	-	-
Chitinases + pNFH	0.96	0.93-1.00	0.94	0.90	-	-	-	-

	ALS vs mimic				ALS vs PLS			
	AUC	CI	Spec.	Sens.	AUC	CI	Spec.	Sens.
Chitotriosidase 1	0.81	0.69-0.92	1.00	0.58	0.81	0.63-0.98	0.83	0.77
Chitinase-3-like protein 1	0.90	0.82-0.99	1.00	0.77	0.59	0.44-0.75	1.00	0.47
Chitinase-3-like protein 2	0.81	0.68-0.95	1.00	0.55	0.87	0.77-0.97	1.00	0.81
Combined chitinases	0.95	0.89-1.00	1.00	0.76	0.88	0.75-1.00	0.83	0.86
pNFH	0.94	0.87-1.00	0.92	0.88	0.88	0.76-0.99	1.00	0.70
Chitinases + pNFH	0.97	0.93-1.00	1.00	0.90	0.93	0.85-1.00	1.00	0.81

Table 3.3: Within-cohort ROC analysis of chitinase proteins in classifying ALS vs non-ALS samples. Chitinase proteins were combined by summation of centered and scaled log abundances for each protein. Performance of the optimal combination of proteins is given (CHIT1 and CHI3L2 for ALS vs PLS, all three proteins for other contrasts). Sensitivity and specificity are given for the Youden index (Youden 1950). ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis; Spec. - specificity; Sens. - sensitivity; pNFH - phosphorylated neurofilament heavy chain.

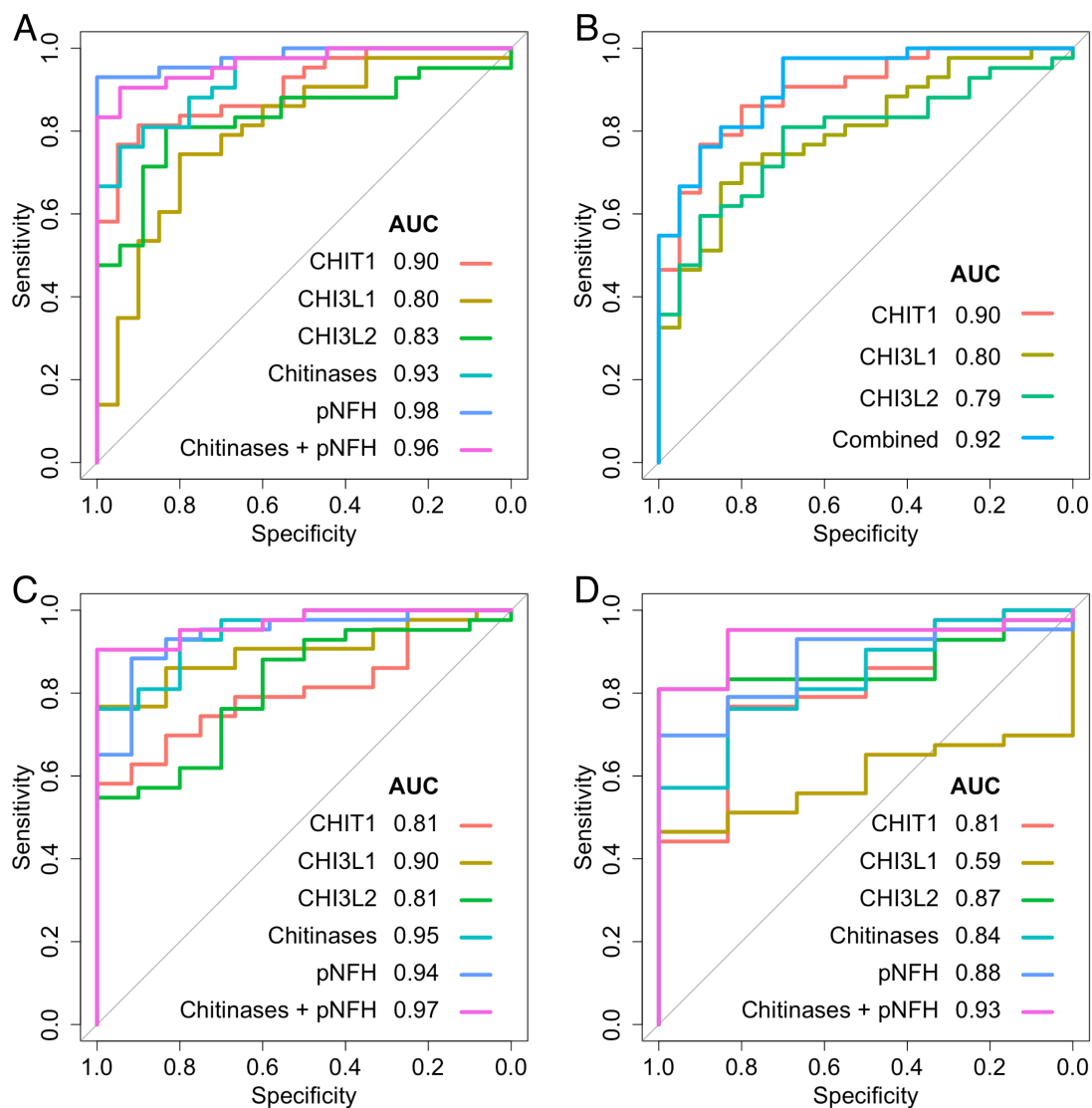


Figure 3.3: Classifier performance of individual and combined chitinase proteins distinguishing ALS from individual control groups by ROC analysis. **A** ALS vs healthy control **B** ALS vs Parkinson's Disease **C** ALS vs mimic conditions **D** ALS vs PLS. Combined analysis performed using summed centered and scaled $\log_2$ abundance of chitinase proteins (the best model is presented including all three chitinase proteins for ALS vs healthy control, ALS vs Parkinson's disease and ALS vs mimic; CHIT1 and CHI3L2 for ALS vs PLS), with and without pNFH. Values for pNFH were not available for Parkinson's disease samples. ALS – amyotrophic lateral sclerosis; PLS – primary lateral sclerosis; CHIT1 – chitotriosidase; CHI3L1 – chitinase-3-like protein 1; CHI3L2 – chitinase-3-like protein 2; pNFH – phosphorylated neurofilament heavy chain.

= 0.009; **Figure 3.3 C** and **Table 3.3**). Performance of the combined chitinase proteins was comparable to pNFH (pNFH AUC 0.94, CI 0.87-1.00; comparison of combined chitinases with pNFH alone $p = 0.912$). The addition of pNFH to the combination of chitinases provided the highest AUC, but did not significantly improve upon the AUC of combined chitinase; (combined chitinases plus pNFH AUC 0.97, CI 0.93-1.00; comparison of combined chitinases plus pNFH versus combined chitinases $p = 0.235$).

Interestingly, though CHIT1 and CHI3L2 showed reasonable performance in differentiating between PLS and ALS (CHIT1 AUC 0.81, CI 0.63-0.98; 0.87, CHI3L2 CI 0.77-0.97; **Figure 3.3 D** and **Table 3.3**), CHI3L1 performed poorly (AUC 0.59, CI 0.44-0.75). A combination of CHIT1 and CHI3L2 increased the AUC non-significantly compared to CHIT1 or CHI3L2 alone (CHIT1 and CHI3L2 combined AUC 0.88, CI 0.75-1.00; comparison of combined chitinases with CHIT1 $p = 0.792$, CHI3L2 $p = 0.600$), and was comparable to pNFH alone (AUC 0.88, CI 0.76-1.00; comparison of combined chitinases and pNFH $p = 0.684$). A combination of CHIT1, CHI3L2 and pNFH provided the highest AUC, though this was not a statistically significant improvement upon CHIT1 and CHI3L2 combined (AUC 0.93, CI 0.85-1.00; comparison of combined chitinases and pNFH $p = 0.072$; **Figure 3.3 D** and **Table 3.3**).

Chitinase proteins also performed well distinguishing between ALS patients and healthy controls (CHIT1 AUC 0.90, CI 0.83-0.98; CHI3L1 AUC 0.80, CI 0.68-0.92; CHI3L2 AUC 0.83, CI 0.73-0.94). Combining chitinases provided a higher AUC, with statistically significant improvement on CHI3L2, but not CHIT1 or CHI3L1 alone (AUC 0.90, CI 0.87-0.99; comparison with CHIT1 alone $p = 0.587$, CHI3L1 $p = 0.064$, CHI3L2 $p = 0.046$). Performance of combined chitinase proteins was similar to pNFH (pNFH AUC 0.98, CI 0.95-1.00; comparison of combined chitinases and pNFH alone $p = 0.157$), and combining chitinase proteins with pNFH worsened performance compared to pNFH alone (combined chitinases plus pNFH AUC 0.96,

CI 0.93-1.00; comparison of combined chitinases plus pNFH and pNFH alone $p = 0.045$; **Figure 3.3 A** and **Table 3.3**).

3.4.4 Cross-sectional linear regression models of chitinase proteins and clinical parameters

Chitinase levels correlates with disease progression rate

Regression models were constructed using samples and clinical data obtained at the first visit of patients with ALS and PLS. Using simple linear regression, CHIT1, CHI3L1 and CHI3L2 levels correlated positively with PR (CHIT1 vs log PR Pearson's $r = 0.56$, $p < 0.001$; CHI3L1 $r = 0.31$, $p = 0.028$; CHI3L2 $r = 0.29$, $p = 0.044$, **3.4 A-C**). There was no correlation with degree of disability as measured by ALSFRS-R for any of the three chitinase proteins (CHIT1 $r = -0.13$, $p = 0.364$; CHI3L1 $r = -0.06$, $p = 0.694$; CHI3L2 $r = -0.12$, $p = 0.431$; **Figure 3.4 D-F**).

Multiple linear regression models controlling for the effect of age, sex, absolute ALSFRS-R, progression rate, and site of onset (bulbar vs spinal) on chitinase protein level demonstrated significant positive correlation with disease progression rate for CHIT1, CHI3L1, and CHI3L2 (CHIT1, $p < 0.001$; gradient = 0.27 log abundance units per unit increase in log disease progression rate; CHI3L1, $p = 0.036$; gradient = 0.05 log abundance units per unit increase in log disease progression rate; CHI3L2, $p = 0.033$; gradient = 0.13 log abundance units per unit increase in log disease progression rate). CHIT1 and CHI3L2 levels were lower in older subjects (CHIT1, $p = 0.025$; gradient = -0.02 log abundance units per year; CHI3L2, $p = 0.032$; gradient = -0.017 log abundance units per year) in cross-sectional analysis of patients with ALS and PLS.

Multiple linear regression models were also constructed to explore the contribution of age and sex across the cohort, modeling each chitinase level separately, with age,

sex, and subject grouping as explanatory variables; no significant association with age or sex was observed.

CHIT1, CHI3L1 and CHI3L2 levels correlate with a marker of axonal damage, pNFH

The neurofilament protein pNFH is a marker of axonal damage and multiple studies have shown significant elevation in ALS CSF that also correlates with PR (see **Chapter 1.3.1**). Levels of CHIT1, CHI3L1 and CHI3L2 all had a moderate positive correlation with pNFH levels (CHIT1 Pearson's $r = 0.62$, $p < 0.001$; CHI3L1 $r = 0.49$, $p < 0.001$; CHI3L2 $r = 0.41$, $p < 0.001$; **Figure 3.4 G-I**).

Additional linear regression analysis was performed to model disease progression rate using first visit protein (CHIT1, CHI3L1, CHI3L2, and pNFH) and the clinical variables age at onset, age at sampling, onset site (bulbar vs spinal), and sex. To provide comparable parameter estimates, levels of CHIT1, CHI3L1, CHI3L2, and pNFH were centered about 0 and scaled to SD 1. A model incorporating all explanatory variables showed significant associations of decreasing age and increasing pNFH with increasing progression rate (model adjusted $R^2 = 0.65$, $p < 0.001$; age of onset gradient = -0.14 log points per month per year, $p < 0.001$; pNFH gradient = 0.69 log points per month per log abundance unit, $p = 0.006$; **Table 3.4**). The most parsimonious model of disease progression rate with highest adjusted R^2 comprised just three explanatory variables: age at symptom onset, CHIT1, and pNFH levels. This demonstrated significant associations of decreasing age, increasing pNFH, and increasing CHIT1 levels with higher disease progression rate (adjusted $R^2 = 0.67$; age at onset, $p < 0.001$; gradient = -0.14 log points per month per year; pNFH $p = 0.005$, gradient = 0.61 log points per month per log abundance unit; CHIT1 $p = 0.045$, gradient = 0.29 points per month per log abundance unit; **Table 3.4**), though it does not take account of clinical parameters that have previously been associated with higher disease progression rate.

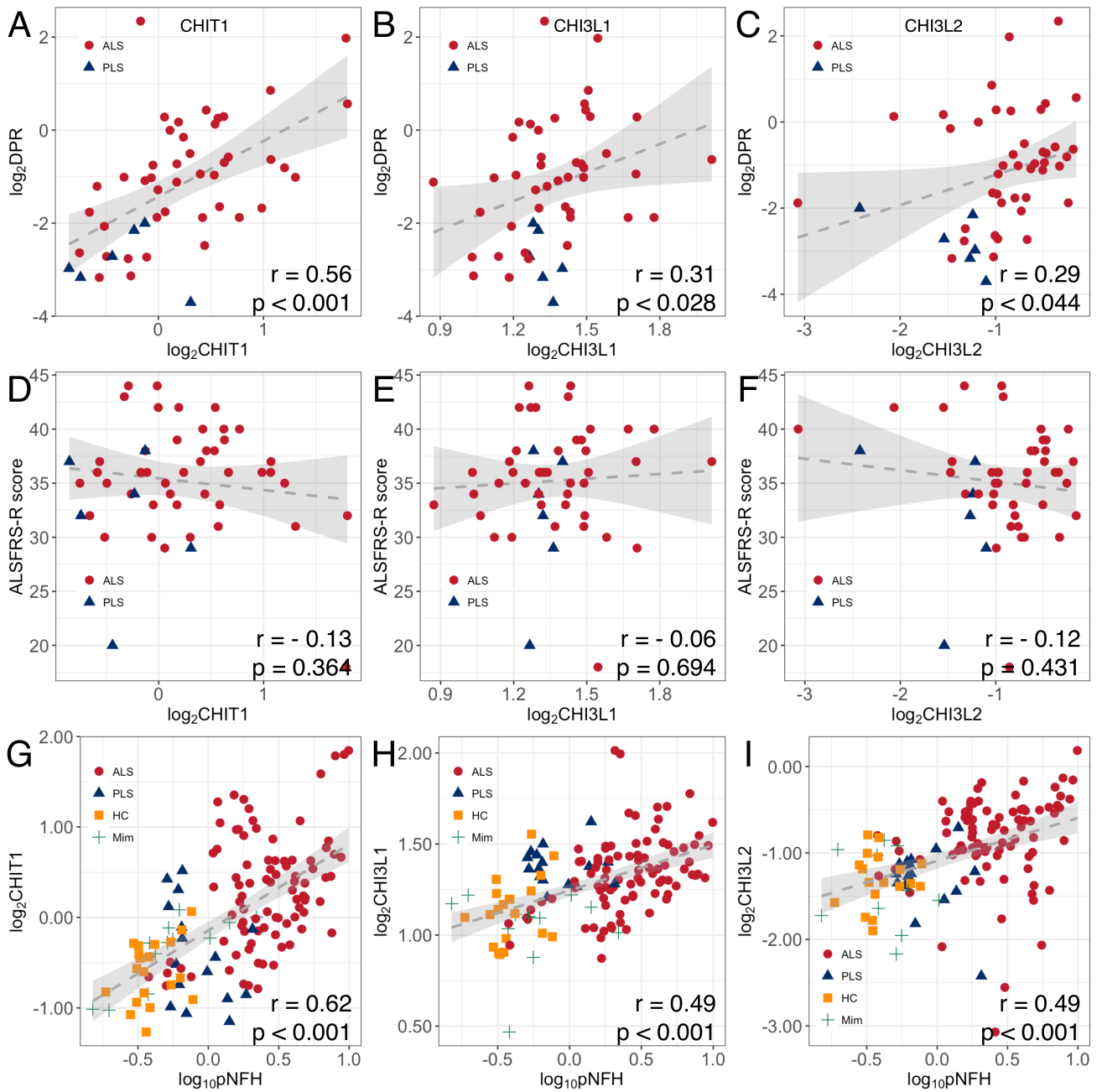


Figure 3.4: Linear regression of clinical parameters. **A–C** correlation between CHIT1, CHI3L1 and CHI3L2 and PR. Pearson’s r and p -value given for simple linear correlation. First visit samples only. **D–F** correlation between CHIT1, CHI3L1 and CHI3L2 and disability as measured by ALSFRS-R score. Pearson’s r and p -value given for simple linear correlation. First visit samples only. **G–I** correlation between CHIT1, CHI3L1 and CHI3L2 and pNFH. Ribbon indicates 95% CI of linear model fit. Pearson’s r and p -value for correlation, all samples. ALS – amyotrophic lateral sclerosis; HC – healthy control; Mim – mimic disorders; PLS – primary lateral sclerosis; PD – Parkinson’s disease; CHIT1 – chitotriosidase; CHI3L1 – chitinase-3-like protein 1; CHI3L2 – chitinase-3-like protein 2; DPR – disease progression rate (ALSFRS-R decline, points per month); ALSFRS-R – revised ALS functional rating scale; pNFH – phosphorylated neurofilament heavy chain.

	Gradient	95% CI	<i>p</i>	Gradient	95% CI	<i>p</i>
Age of onset	-0.13	-0.16 to -0.06	<0.001	-0.14	-0.22 to -0.06	<0.001
pNFH	0.61	0.19 - 1.04	0.006	0.69	0.21 - 1.17	0.006
CHIT1	0.29	0.01 - 0.57	0.045	0.29	-0.04 - 0.61	0.085
CHI3L1	-	-	-	-0.10	-0.40 - 0.19	0.487
CHI3L2	-	-	-	0.07	-0.16 - 0.30	0.537
Sex (male)	-	-	-	0.25	-0.79 - 0.29	0.360
Site of onset (limb)	-	-	-	0.00	-0.61 - 0.61	0.999
Adjusted R^2	0.67	-	<0.001	0.65	-	<0.001

Table 3.4: Multiple Linear Regression Modeling of Disease Progression Rate Using Chitinase Proteins, pNFH, and Clinical Parameters. ALS – amyotrophic lateral sclerosis; PLS – primary lateral sclerosis; CI - confidence interval.

3.4.5 Longitudinal analysis of chitinase proteins

A linear mixed model of CSF CHIT1, CHI3L1 and CHI3L2 was fitted to longitudinal values, beginning at the baseline visit, with patients stratified by level at baseline visit in relation to median value. Longitudinal sample numbers for ALS and PLS are summarized in **Table 3.1**. There was a small increase in CHI3L1 levels over time in those with low levels at onset (gradient = 0.005 log abundance units per month, $p = 0.001$). The gradient did not differ significantly from zero in those with high initial CHI3L1 or for either stratum of CHIT1 or CHI3L2 (**Figure 3.5**).

3.4.6 Survival analysis

A Cox proportional hazards model including first visit ALSFRS-R, age at onset, sex, onset site (bulbar vs spinal), and CHIT1, CHI3L1, and CHI3L2 levels in all MND patients showed association of CHIT1 level and age with survival (CHIT1 hazard ratio (HR) 2.84, $p = 0.009$, 95% CI 1.24–5.75; CHI3L1 HR = 1.23, $p = 0.55$, 95% CI 0.64–2.33; CHI3L2 HR = 1.08, $p = 0.93$, 95% CI 0.20–7.13; age HR = 1.00015, $p = 0.008$, 95% CI 1.00004–1.00027; 3.5). The addition of pNFH level to the model left only pNFH as an independent prognostic factor (pNFH HR = 1.26, $p = 0.019$, 95%

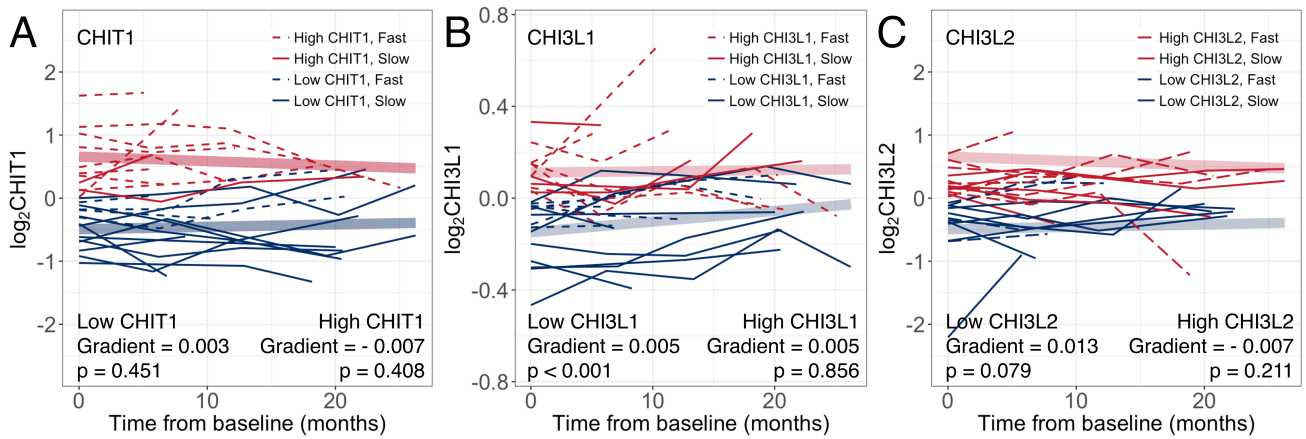


Figure 3.5: Longitudinal stability of CHIT1, CHI3L1, and CHI3L2. Gradient estimated using a random slope, random intercept linear mixed model, stratified by initial analyte level. Ribbon indicates mean predicted longitudinal change by stratum. Gradient in $\log_2$ units change per month; p value denotes likelihood that gradient differs from zero. Patients stratified by disease progression rate at last visit. CHIT1 - chitotriosidase; CHI3L1 - chitinase-3-like protein 1; CHI3L2 - chitinase-3-like protein 2.

CI 1.04–1.53; CHIT1 HR = 1.90, $p = 0.137$, 95% CI 0.82–4.44; CHI3L1 HR = 1.37, $p = 0.765$, 95% CI 0.18–10.60; CHI3L2 HR = 0.88, $p = 0.728$, 95% CI 0.43–1.81). Data for several factors known to influence survival in ALS including nutritional status, *C9orf72* genotype and respiratory function (forced vital capacity (FVC), for example) were not available, so caution is needed in drawing firm conclusions around an independent role for CHIT1 levels at present (Westeneng et al. 2018).

Kaplan-Meier survival curves showed apparent separation of cumulative survival of patients when stratified by level of CHIT1, CHI3L1, or CHI3L2, though no significant association between stratum and survival was observed (log-rank test, CHIT1 $p = 0.175$, CHI3L1 $p = 0.082$, CHI3L2 $p = 0.149$; 3.6).

3.4.7 Validation of previously identified candidate markers

Of the previously identified candidate proteins differentiating ALS patients from healthy or disease controls, 6 proteins were identified in the current study: Cystatin C (CST3; Ranganathan et al. 2005; Pasinetti et al. 2006; Ranganathan et al.

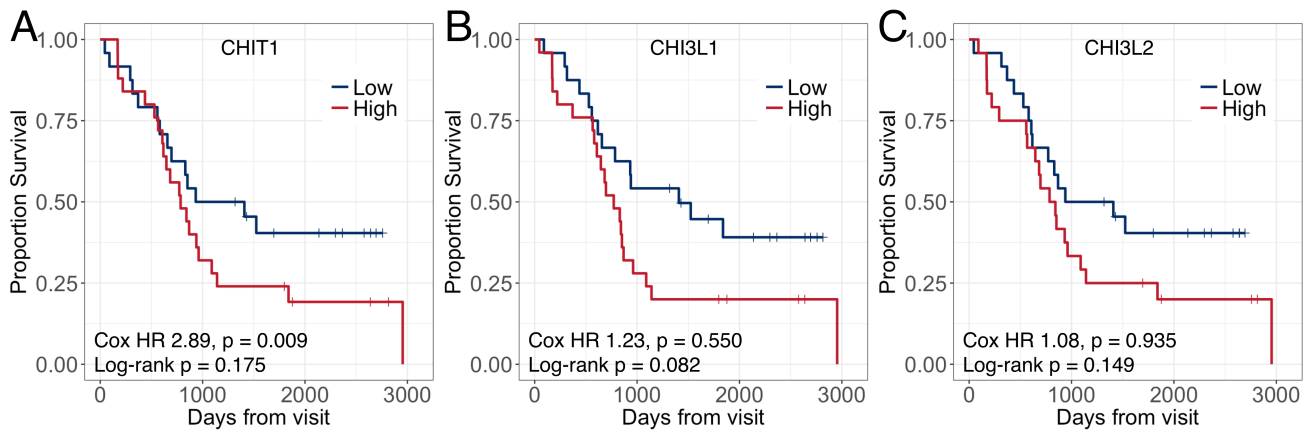


Figure 3.6: Survival analysis in ALS and PLS patients from time of first sampling, stratified by relative abundance of CHIT1, CHI3L1, and CHI3L2 at first visit. Cox proportional hazards model hazard ratio and p -value listed along with p -value for log-rank test. CHIT1 - chitotriosidase; CHI3L1 - chitinase-3-like protein 1; CHI3L2 - chitinase-3-like protein 2; HR - hazard ratio.

	HR	95% CI	p	HR	95% CI	p
Age of onset	1.00	1.00004-1.00027	0.008	1.00	1.00000-1.00019	0.149
Sex (male)	1.32	0.53-2.57	0.488	0.96	0.41-2.21	0.916
ALSFRS-R at baseline	1.08	0.97-1.18	0.141	1.04	0.95-1.15	0.359
Site of onset (limb)	0.74	0.30-2.11	0.531	0.67	0.27-1.70	0.396
CHIT1	2.84	1.24-5.75	0.009	1.90	0.82-4.40	0.137
CHI3L1	1.23	0.64-2.33	0.549	1.37	0.18-10.60	0.765
CHI3L2	1.08	0.20-7.13	0.935	0.89	0.43-1.81	0.728
pNFH	-	-	-	1.26	1.04-1.53	0.019

Table 3.5: Cox Proportional Hazards Survival Models for Patients With ALS and PLS, Based on First Visit Analyte Value, With and Without Inclusion of pNFH. ALS – amyotrophic lateral sclerosis; PLS – primary lateral sclerosis; HR - hazard ratio; CI - confidence interval.

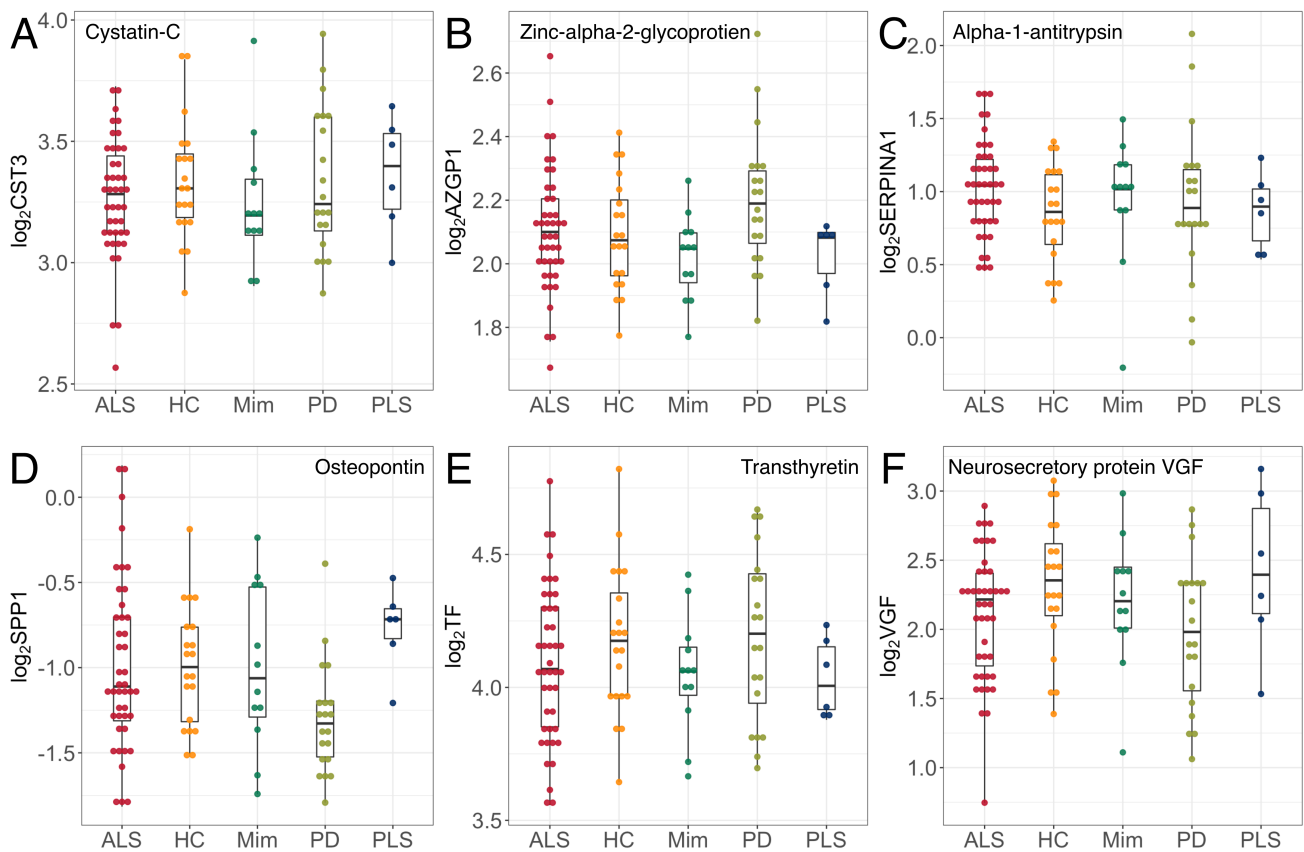


Figure 3.7: Relative abundance of candidate protein biomarkers identified in previous proteomic studies using data from the current study. CST3 - Cystatin-C; AZGP1 - Zinc-alpha-2-glycoprotein; SERPINA1 - Alpha-1-antitrypsin; SPP1 - Osteopontin; TTR - Transthyretin; VGF - Neurosecretory protein VGF.

2007; Ryberg et al. 2010; Collins et al. 2015), Zinc-alpha-2-glycoprotein (AZGP1; Brettschneider et al. 2008; Chen et al. 2016b), Alpha-1-antitrypsin (SERPINA1; Brettschneider et al. 2008; Neuhoﬀ et al. 2012; Chen et al. 2016b), Osteopontin (SPP1; Varghese et al. 2013), Transthyretin (TTR; Ranganathan et al. 2005; Ranganathan et al. 2007; Ryberg et al. 2010) and Neurosecretory protein VGF (VGF Pasinetti et al. 2006; Neuhoﬀ et al. 2012). Across all comparisons, only Osteopontin diﬀered between ALS and Parkinson’s disease patients (ALS mean = -0.99 log abundance units, Parkinson’s disease mean = -1.30 log abundance units, $p = 0.004$). Following adjustment for multiple comparisons, no signiﬁcant diﬀerences were observed for any protein across any comparison.

3.4.8 Discussion

The work of this chapter describes the proteomic analysis of individual CSF samples from a large longitudinal cohort of patients with ALS and relevant control conditions including neurodegenerative disease controls, ALS-mimicking conditions and PLS. The simple preanalytical methods, using a high-throughput tryptic digestion, resulted in reliable quantification across the cohort.

The main findings were elevation of a set of three related chitinase proteins, CHIT1, CHI3L1 and CHI3L2. Elevation of all three proteins in the CSF of ALS patients compared with healthy controls was previously demonstrated in a small proteomic study, though a within-cohort validation by ELISA only demonstrated elevation of CHIT1 (Varghese et al. 2013). Separately, elevation of CHIT1 in ALS has been confirmed by ELISA (Steinacker et al. 2018; Chen et al. 2016a). Elevation of CHI3L2 was also demonstrated in a shotgun proteomic study, though this has not been validated by immunoassay (Collins et al. 2015).

In addition, the work described here explored the classifier performance of the chitinase proteins as measured by proteomics and found it to be similar to, though no better than, pNFH. Interestingly, elevation of CHIT1 and CHI3L2, but not CHI3L1 was observed in ALS compared to PLS. This could suggest differences in the glial response between the two conditions and suggests that biochemically distinguishing between the two conditions might be possible. Notably, the number of PLS samples used in the analysis was small and the latency from symptom onset to sampling was comparatively high. Confirmation of the findings in a prospectively-acquired cohort during the initial diagnostic phase is essential prior to clinical translation.

Correlation of the chitinase proteins with clinical parameters indicated a relationship with the aggressiveness of the disease rather than the stage of the disease, similar to the neurofilament proteins. Longitudinal analysis indicated little change of the

proteins over time, with the exception of CHI3L1 which increased gradually in those with initially low levels. This implies that glial activity does not rise rapidly through the course of the disease.

Survival analysis showed a relationship of higher levels of CHIT1 with shortened survival, though this association was less robust than that of pNFH. Again, validation of this in cohorts incorporating the full spectrum of survival-associated parameters of ALS is required prior to translation of these findings to the pharmaceutical trial or clinical arena.

All three of the chitinase proteins are known to be released by macrophages, with roles in the immune response and tissue remodelling, though their origin within the central nervous system is less well-established (Di Rosa et al. 2016). Limited work suggests that CHIT1 is secreted by microglia *in vitro* and colocalises with the microglial marker IBA-1 in the spinal cord of ALS patients *post mortem* (Varghese et al. 2013; Steinacker et al. 2018). The origin of CHI3L1 and CHI3L2 has not been studied in ALS, though a reanalysis of RNA expression data from motor cortex of ALS patients suggests that expression of these two proteins is increased in ALS (Sanfilippo et al. 2017).

Although the data presented in this chapter cannot exclude overspill from plasma as the cause of elevation of these proteins, current understanding of the origin of these proteins would suggest that their levels are likely to represent glial activity within the CNS. The contribution of glial mechanisms to the pathogenesis of ALS is controversial, but there is evidence that they may at least represent a modifier of severity that is amenable to treatment. As a consequence, glial pathways are being actively pursued in clinical trials. Chitinase proteins represent a promising means to measure these mechanisms and the effect of potential glial-targeting treatments. Further work is necessary to validate these findings and to understand the cellular and CNS regional origin of these proteins before firm conclusions can be drawn. Correlation of CSF measurement of chitinase proteins with *in vivo* measures of glial

activity, such as microglial PET, and *post mortem* analysis of protein expression would go a significant way to achieving this.

Although six additional candidate biomarkers identified in previous proteomic studies of CSF in ALS were quantified here, only elevation of osteopontin in relation to Parkinson's disease was detected (and even this was lost following correction for multiple comparisons). This finding is also contradictory to previous observations in Parkinson's disease, in which elevations of CSF osteopontin have been observed when compared to neurological disease controls (Maetzler et al. 2009).

There are multiple possible explanations for these discrepancies, including differences in sample handling (Cystatin C, for example, is susceptible to degradation when stored at -20°C; Barschke et al. 2017), differences in the cohort selected; pre-analytical steps such as abundant protein depletion, prefractionation or labelling; differing MS methodology (previous studies have used separation techniques that might lead to overrepresentation of high molecular weight proteins); and differences in the normalisation strategy and statistical analysis (Silberring et al. 2010).

4. Validation of Chitinase Biomarkers in ALS

4.1 Hypothesis and Aims of This Chapter

4.1.1 Hypothesis

Chitinase proteins are elevated in the CSF of ALS patients and may be useful in diagnosis, prognostication and reflect underlying glial activity, making them useful target engagement markers and predictors of response to glial-targeted therapies for ALS.

4.1.2 Aims

1. Validate the findings of elevated CSF chitinase proteins in ALS using an orthogonal method (ELISA)
2. Examine associations of the chitinase proteins with clinical parameters:
 - i. Rate of disease progression
 - ii. Upper motor neuron burden
 - iii. Cognitive dysfunction
 - iv. Longitudinal change
 - iv. Survival

3. Examine associations of the chitinase proteins with biochemical measures:
 - i. Axonal degeneration: pNFH
 - ii. Nonspecific inflammatory response: CRP
 - iii. Chitotriosidase activity
 - iv. Serum chitinase protein levels
4. Examine associations of the chitinase proteins with structural imaging parameters

4.2 Introduction

In addition to the work outlined in **Chapter 3**, elevations in the chitinase proteins in the CSF of patients with ALS has now been demonstrated in several studies using MS-based proteomics and ELISA (Varghese et al. 2013; Collins et al. 2015; Steinacker et al. 2018; Chen et al. 2016a). These studies have predominantly focused on CHIT1, confirming its elevation in ALS compared to healthy controls and those with other neurodegenerative diseases (with the exception of CJD) and demonstrating good classifier performance distinguishing ALS patients from healthy controls, albeit with diminished accuracy compared to CSF neurofilament levels (Chen et al. 2016a; Steinacker et al. 2018).

Examination of *post mortem* spinal cord tissue from ALS patients using double-labelled immunofluorescence has identified microglia of the corticospinal tract as the likely cause of elevated CSF CHIT1, through localisation of CHIT1 immunofluorescence in the somata of IBA-1 positive (macrophage lineage) cells (Steinacker et al. 2018). Notably, however, examination of affected brain regions and co-immunofluorescence of markers for neurons and other glial cells has not been performed.

CHI3L1 and CHI3L2 have been less well-studied in ALS, with only one proteomic study demonstrating elevation of CSF levels of both proteins (with ELISA validation

in the same study confirming elevation of just CHI3L2; Varghese et al. 2013) and another study demonstrating elevation of just CHI3L2 without additional validation (Collins et al. 2015). Evidence regarding the CNS cellular origin of CHI3L1 and CHI3L2 is also less clear than for CHIT1, with both proteins expressed in microglia and astrocytes in white matter of multiple sclerosis patients, but CHI3L1 expressed exclusively by astrocytes in frontal cortex tissue of patients with Alzheimer’s disease (Hinsinger et al. 2015; Querol-Vilaseca et al. 2017). No studies yet examine the expression of these two proteins in ALS *post mortem* tissue.

This chapter aims to validate the findings of the **Chapter 3** with respect to elevations of CSF CHIT1, CHI3L1 and CHI3L2 using commercially-available immunoassays to provide absolute abundance measurements, both within the cohort studied and in an independently-acquired cohort. The accuracy of classification, relationship with clinical, biochemical and imaging parameters, and longitudinal stability will also be explored.

4.3 Materials and methods

4.3.1 Participants and sampling

This study included 82 patients with ALS, 11 patients with PLS, 12 patients with mimic conditions (including multifocal motor neuropathy with conduction block, Kennedy’s disease, cervical spondylotic myeloradiculopathy and motor presentation of multiple sclerosis), 25 healthy control subjects and 5 subjects carrying the *C9orf72* hexanucleotide repeat expansion but not manifesting symptoms. Of this cohort, 39 ALS patients, 5 healthy controls, 4 PLS patients and 5 asymptomatic gene carriers had not been analysed using proteomics. Subject recruitment, CSF sampling, sample processing and clinical characterisation was completed as described in **Section 2.3**. Demographic and clinical characteristics are displayed in **Table 4.1**.

	ALS	PLS	Mimic conditions	Healthy control	Asymptomatic gene carriers	<i>p</i>
n, first visit	82	10	12	25	5	-
n, second visit	33	5	-	-	-	-
n, third visit	19	3	-	-	-	-
n, fourth visit	10	2	-	-	-	-
n, fifth visit	2	-	-	-	-	-
% male	76.8	40.0	91.7	48.0	20.0	0.013*
Age at sampling, mean $\pm$ SD (years)	59.9 $\pm$ 10.7	66.4 $\pm$ 9.7	57.7 $\pm$ 15.5	56.5 $\pm$ 9.2	39.7 $\pm$ 7.1	0.011 ⁺
Age at onset, mean $\pm$ SD (years)	56.8 $\pm$ 11.2	44.0 $\pm$ 26.0	-	-	-	0.010
% bulbar onset	26.8	10	-	-	-	0.438
Disease progression rate median [IQR]	0.45 [0.23-0.60]	0.12 [0.08-0.17]	-	-	-	0.040
Deaths, n (%)	44 (54)	0	-	-	-	< 0.001

Table 4.1: Demographic data. *p*-values for χ^2 test (categorical data) or Kruskal-Wallis *H* test (continuous data; or Mann-Whitney *U* test for two-group comparisons). **Post hoc* pairwise Fisher exact tests with Benjamini-Hochberg FDR correction indicated significantly more male subjects in ALS and mimics compared to other groups. ⁺Dunn's *post hoc* with Benjamini-Hochberg FDR correction *p* < 0.05 for asymptomatic gene carriers versus other groups. ALS – amyotrophic lateral sclerosis; PLS – primary lateral sclerosis.

4.3.2 ELISA and CHIT1 activity assay

All ELISAs and activity assays were performed in duplicate. Plates were read using a FLUOstar Omega plate reader (BMG LABTECH, UK). Standard curves were fitted with 4-parameter logistic regression using MARS data analysis software (BMG LABTECH, UK). CSF samples were thawed on ice prior to measurement. Chitinase protein ELISAs were tested for intra-assay CV, interassay CV, LLOQ, dilutional linearity and parallelism as outlined in Andreasson et al. (2015).

CHIT1

CHIT1 level was measured using a commercially available ELISA (CircuLex™ human Chitotriosidase ELISA, MBL, UK) according to the manufacturer's instructions. CSF samples were diluted to 1 in 20 with dilution buffer and vortexed before adding to the plate. Samples above the top standard were subjected to repeat assay at 1 in 100 dilution, those below the bottom standard were repeated at 1 in 5 dilution. Serum samples were assayed at 1 in 25 dilution. 100µL of diluted standards and samples were added to the plate in duplicate and incubated for 1 hour at room temperature with 400rpm shaking. The plate was washed four times with 350µL wash buffer. 100µL horseradish peroxidase (HRP)-conjugated anti-human CHIT1 antibody was added and incubated for 1 hour at room temperature with 400rpm shaking. The plate was washed four times with 350µL wash buffer. 100µL substrate reagent was added and incubated at room temperature for 20 minutes with 400rpm shaking. 100µL stop solution was added and plates were read immediately at dual wavelengths (450-540nm).

The median intra-assay CV was 2.58% (IQR 1.22-4.80%), interassay CV median 15.04% (IQR 14.38-15.84%). The LLOQ was 70.95 pg/mL. CSF CHIT1 levels showed dilutional linearity and parallelism in the linear range of the standard curve.

CHI3L1

CHI3L1 level was measured using a commercially available ELISA (Human Chitinase 3-like 1 Quantikine ELISA kit, R&D systems, UK) according to the manufacturer's instructions. CSF samples were diluted to 1 in 250 with dilution buffer and vortexed before adding to the plate. Serum samples were assayed at 1 in 50 dilution. 100µL of assay buffer was added to each well. 50µL of diluted standards and samples were added to the plate in duplicate and incubated for 2 hours at room temperature with 400rpm shaking. The plate was washed four times with 400µL wash buffer. 200µL HRP-conjugated anti-human CHI3L1 antibody was added and incubated for 2 hours at room temperature with 400rpm shaking. The plate was washed four times with 400µL wash buffer. 200µL substrate reagent was added and incubated at room temperature for 30 minutes with 400rpm shaking. 50µL stop solution was added and plates were read immediately at dual wavelengths (450-540nm).

The median intra-assay CV was 1.86% (IQR 0.90-3.61%), interassay CV median 2.80% (IQR 2.70-5.90%). The LLOQ was 100.02 pg/mL. CSF CHI3L1 levels showed dilutional linearity and parallelism in the linear range of the standard curve.

CHI3L2

CHI3L2 level was measured using a commercially available ELISA (CircuLex™ human YKL-39 ELISA, MBL, UK) according to the manufacturer's instructions. CSF samples were diluted to 1 in 40 with dilution buffer and vortexed before adding to the plate. Serum samples were assayed at 1 in 4 dilution. 100µL of diluted standards and samples were added to the plate in duplicate and incubated for 1 hour at room temperature with 400rpm shaking. The plate was washed four times with 350µL wash buffer. 100µL HRP-conjugated anti-human YKL-39 antibody was added and incubated for 1 hour at room temperature with 400rpm shaking. The plate was washed four times with 350µL wash buffer. 100µL substrate reagent

was added and incubated at room temperature for 20 minutes with 400rpm shaking. 100µL stop solution was added and plates were read immediately at dual wavelengths (450-540nm).

The median intra-assay CV was 2.61% (IQR 1.27-5.20%), interassay CV median 5.07% (IQR 4.05-5.10%). The LLOQ was 58.34 pg/mL. CSF CHI3L2 levels showed dilutional linearity and parallelism in the linear range of the standard curve.

pNFH

pNFH level was measured as described in **Section 3.3.4**.

CRP

CRP level was measured using a commercially available ELISA (Abcam[®] CRP human ELISA kit, Abcam[®], UK) according to the manufacturer's instructions. CSF samples were diluted to 1 in 2 with sample diluent and vortexed before adding to the plate. Samples below the bottom standard were re-assayed neat, those above the top standard were subjected to repeat assay at a 1 in 4 dilution. 50µL of diluted standards and samples were added to the plate in duplicate and incubated for 2 hour at room temperature with 400rpm shaking. The plate was washed five times with 200µL wash buffer. 50µL biotinylated anti-human CRP antibody was added and incubated for 30 minutes at room temperature with 400rpm shaking. The plate was washed five times with 200µL wash buffer. 50µL streptavidin-peroxidase conjugate was added and incubated for 30 minutes at room temperature with 400rpm shaking. The plate was washed five times with 200µL wash buffer. 50µL chromogen substrate reagent was added and incubated at room temperature for 10 minutes with 400rpm shaking. 50µL stop solution was added and plates were read immediately at dual wavelengths (450-570nm).

The median intra-assay CV was 4.67% (IQR 2.12-8.00%), interassay CV median 17.22% (IQR 11.81-22.64%). The LLOQ was 175.81 pg/mL.

4.3.3 CHIT1 activity measurement

CHIT1 activity was measured using a commercially available fluorometric assay (CycLex[®] Chitotriosidase Fluorimetric Assay Kit, MBL, UK) according to the manufacturer's instructions. CSF samples were diluted to 1 in 5 with inhibition or assay buffer. 5µL of diluted standards (4-Methylumbelliferone), samples in inhibition and assay buffer and positive controls (recombinant CHIT1) were added to the plate in duplicate. 95µL reaction buffer was added. Fluorescence intensity was read using a FLUOstar Omega plate reader (BMG LABTECH, UK) with excitation at 320nm and emission at 445nm, recording in kinetic mode for 41 minutes. CHIT1 activity was calculated using the change in blank-corrected fluorescence units at 10- and 40-minute time-points, subtracting non-specific substrate hydrolysis.

The median intra-assay CV was 2.38% (IQR 1.51-5.61%), interassay CV 13.90%.

4.3.4 Statistical analysis

Statistical analysis was performed as described in **Section 2.4**. Protein levels between groups were compared using a pairwise Mann-Whitney *U* test with Benjamini-Hochberg multiple comparison correction. Correlation with clinical parameters was calculated using Pearson's correlation coefficient and associations between predictor and outcome variables tested using Student's *t*-test. Cox regression analysis of survival was performed incorporating known survival covariates and univariate analysis using the log-rank test, again using data from first visit samples. Longitudinal data analysis was performed using a random intercept, random slope linear mixed model.

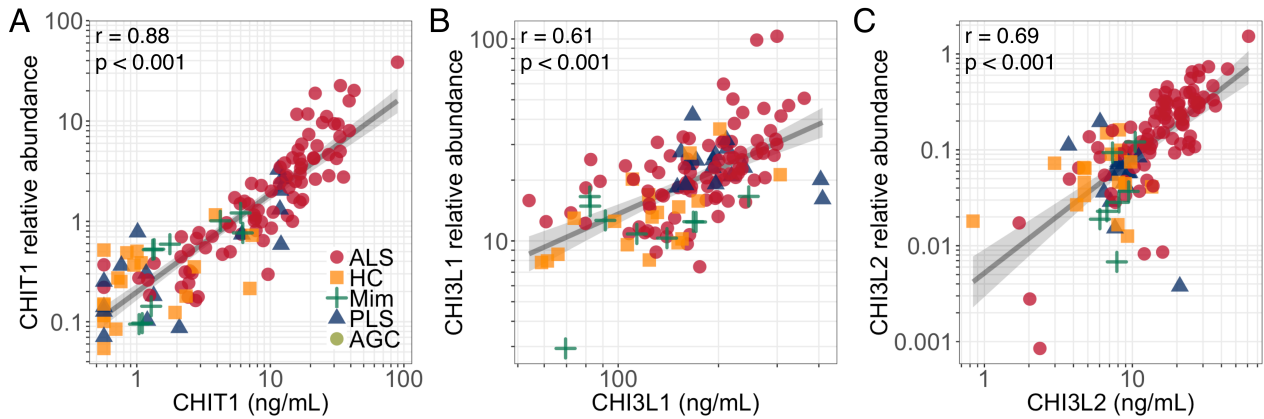


Figure 4.1: Correlation of proteomic and ELISA measurement of **A** CHIT1 **B** CHI3L1 and **C** CHI3L2

4.4 Results

4.4.1 Demographic data

Subject age differed by group ($p = 0.011$). *Post hoc* testing demonstrated only significant differences between the asymptomatic gene carriers and other groups, the asymptomatic gene carriers being younger (ALS $p = 0.013$, PLS $p = 0.010$, mimics $p = 0.047$, healthy controls $p = 0.043$). Gender was also different between groups ($p = 0.013$). There were significantly more male subjects in the ALS and mimic groups compared to other groups on *post hoc* testing (ALS vs PLS $p = 0.037$, healthy controls $p = 0.037$, asymptomatic gene carriers $p = 0.037$; mimic vs PLS $p = 0.037$, healthy controls $p = 0.037$, asymptomatic gene carriers $p = 0.037$; **Table 4.1**).

4.4.2 Correlation of ELISA and proteomic data

ELISA and proteomic measurement of each chitinase protein was correlated (CHIT1 Pearson's $r = 0.88$, $p < 0.001$; CHI3L1 $r = 0.61$, $p < 0.001$; CHI3L2 $r = 0.69$, $p < 0.001$ **Figure 4.1**).

4.4.3 Cross-sectional chitinase protein levels

In accordance with findings in the proteomic analysis described in **Chapter 3**, the CSF levels of CHIT1, CHI3L1 and CHI3L2 were elevated in the first-visit CSF samples of ALS patients compared to healthy controls (FDR-adjusted $p < 0.001$ for all chitinase proteins), mimics (CHIT1 $p < 0.001$, CHI3L1 $p = 0.017$, CHI3L2 $p < 0.001$) and asymptomatic gene carriers (CHIT1 $p = 0.003$, CHI3L1 $p < 0.001$, CHI3L2 $p = 0.002$). Levels of CHIT1 and CHI3L2 were elevated in ALS compared to PLS, though levels of CHI3L1 were similar in both ALS and PLS (CHIT1 $p = 0.021$, CHI3L1 $p = 0.417$, CHI3L2 $p < 0.001$; **Figure 4.2A-C**).

Levels of all three chitinase proteins were similar between healthy control subjects and asymptomatic gene carriers (CHIT1 $p = 0.844$, CHI3L1 $p = 0.0657$, CHI3L2 $p = 0.578$).

To demonstrate alterations of CSF chitinase levels in ALS patients in an independent cohort, analysis of samples not included in the earlier proteomic analysis was performed. This showed significant elevation of CHIT1 in ALS samples compared with healthy controls (FDR-adjusted $p < 0.001$) and asymptomatic gene carriers ($p < 0.001$) but no significant difference between ALS or PLS ($p = 0.112$). CHI3L1 was elevated in ALS compared with healthy controls ($p < 0.001$) asymptomatic gene carriers ($p < 0.001$) and PLS patients ($p = 0.015$), and CHI3L2 was also elevated in ALS compared with healthy controls ($p < 0.001$), asymptomatic gene carriers ($p < 0.001$) and PLS patients ($p = 0.045$; **Figure 4.2**).

ELISA data for all subjects (i.e. regardless of whether previously analysed by proteomics) was included in subsequent analyses.

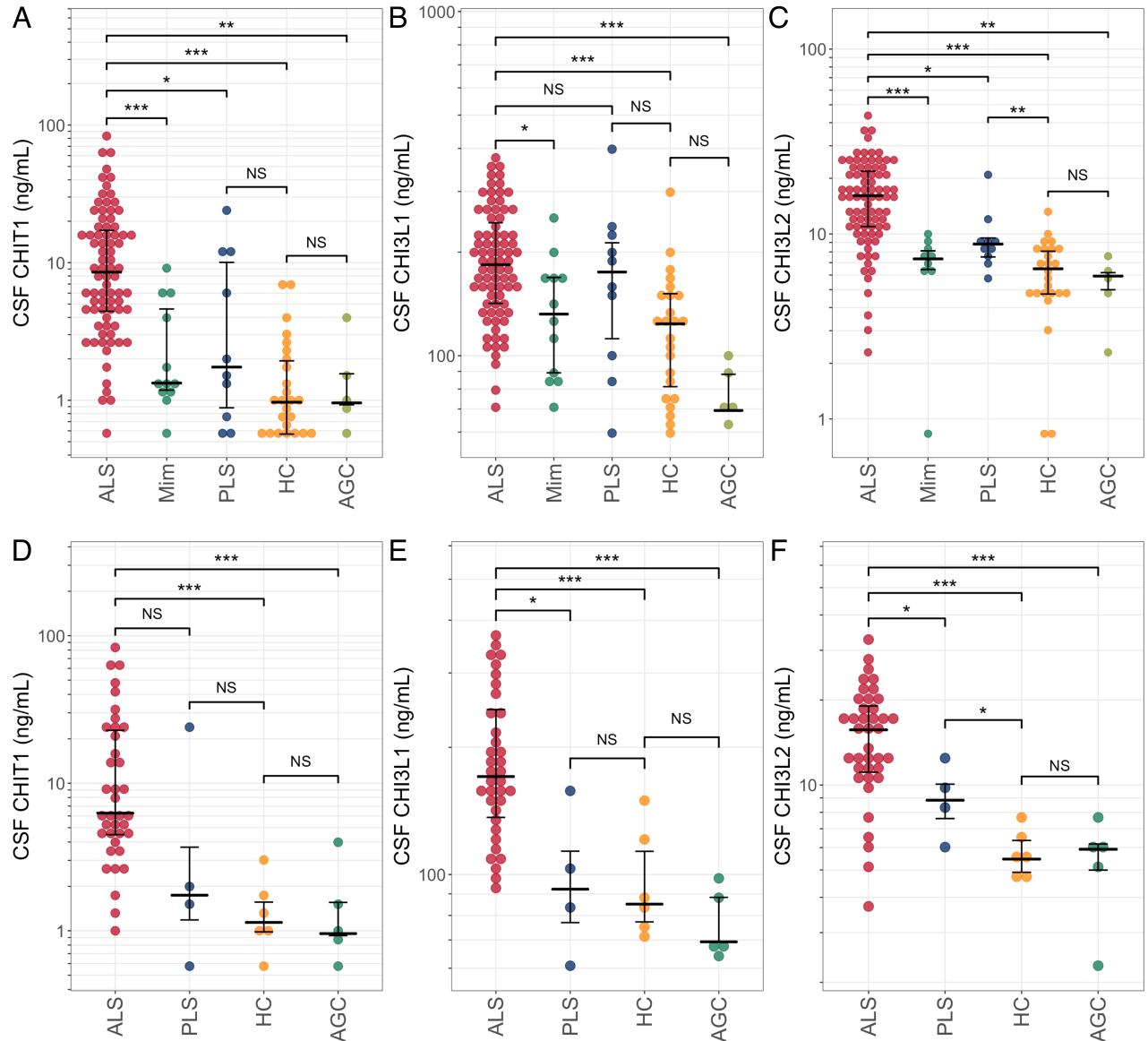


Figure 4.2: Cross-sectional CSF levels of **A, D** CHIT1, **B, E** CHI3L1 and **C, F** CHI3L2 in the entire cohort (**A-C**) and the cohort not included in proteomic analysis (**D-F**). *FDR-adjusted $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, NS $p > 0.05$. ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis; Mim - mimic disorders; HC - healthy control; AGC - asymptomatic gene carrier; CHIT1 - Chitotriosidase-1; CHI3L1 - Chitinase-3-like protein 1; CHI3L2 - Chitinase-3-like protein 2.

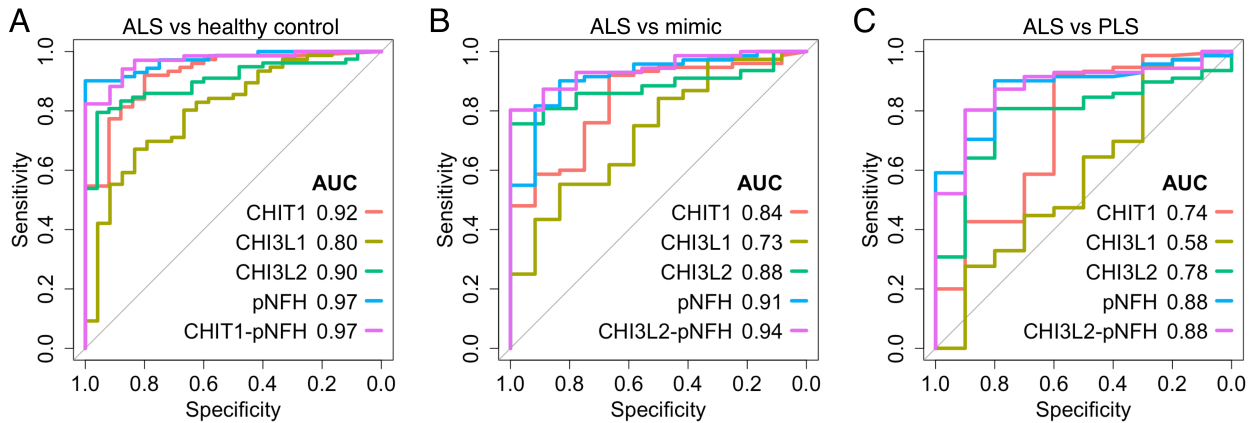


Figure 4.3: Performance of chitinase proteins and pNFH singly and in combination distinguishing CSF of ALS patients from healthy controls **A**, patients with mimic conditions **B** and patients with PLS **C**. ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis; CHIT1 - Chitotriosidase-1; CHI3L1 - Chitinase-3-like protein 1; CHI3L2 - Chitinase-3-like protein 2.

4.4.4 Classifier performance

Chitinase protein diagnostic performance (including sensitivity and specificity at the maximal Youden index; Youden 1950) is outlined in **Figure 4.3** and **Table 4.2**. All chitinase proteins performed relatively well in distinguishing ALS from healthy controls and mimics, with CHI3L1 performing least well; individual chitinase proteins did not outperform pNFH alone, nor did the optimal combination of chitinase proteins and pNFH (ALS vs healthy controls: CHIT1 AUC 0.92 [95% CI 0.86-0.98], CHI3L1 0.80 [0.70-0.91], CHI3L2 0.90 [0.84-0.96], pNFH 0.97 [0.94-1.00], CHIT1 plus pNFH 0.97 [0.94-1.00], difference between AUC of pNFH and pNFH plus CHIT1 $p = 0.841$; mimic CHIT1 0.84 [0.72-0.95], CHI3L1 0.73 [0.58-0.88], CHI3L2 0.88 [0.81-0.95], pNFH 0.91 [0.83-0.99], CHI3L2 plus pNFH 0.94 [0.88-1.00], difference between AUC of pNFH and pNFH plus CHI3L2 $p = 0.608$).

Performance distinguishing ALS from PLS was worse, particularly using CHI3L1 (CHIT1 0.74 [0.55-0.93], CHI3L1 0.58 [0.36-0.79], CHI3L2 0.78 [0.65-0.91], pNFH 0.88 [0.78-0.97], CHI3L2 plus pNFH 0.88 [0.78-0.98], difference between AUC of pNFH and pNFH plus CHI3L2 $p = 0.973$).

	ALS vs healthy control				ALS vs Mimic			
	AUC	CI	Spec.	Sens.	AUC	CI	Spec.	Sens.
Chitotriosidase 1	0.92	0.86-0.98	0.8	0.92	0.84	0.72-0.95	0.67	0.92
Chitinase-3-like protein 1	0.80	0.70-0.91	0.83	0.67	0.73	0.58-0.88	0.83	0.55
Chitinase-3-like protein 2	0.90	0.84-0.96	0.96	0.79	0.88	0.81-0.95	1.00	0.76
pNFH	0.97	0.94-1.00	1.00	0.90	0.91	0.83-0.99	0.83	0.90
Optimal Combination	0.97*	0.94-1.00*	1.00*	0.82*	0.94 ⁺	0.88-1.00 ⁺	1.00 ⁺	0.80 ⁺

	ALS vs PLS			
	AUC	CI	Spec.	Sens.
Chitotriosidase 1	0.74	0.55-0.93	0.60	0.92
Chitinase-3-like protein 1	0.58	0.36-0.79	0.3	0.95
Chitinase-3-like protein 2	0.78	0.65-0.91	0.80	0.81
pNFH	0.88	0.78-0.97	0.80	0.90
Optimal Combination	0.88 ⁺	0.78-0.98 ⁺	0.9 ⁺	0.80 ⁺

Table 4.2: Within-cohort ROC analysis of chitinase proteins in classifying ALS vs non-ALS samples. Protein levels were combined by summation of log abundances. Performance of the optimal combination of proteins is given. Sensitivity and specificity are given for the maximal Youden index. *Optimal combination CHIT1 and pNFH, ⁺Optimal combination CHI3L2 and pNFH. ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis; Spec. - specificity; Sens. - sensitivity; pNFH - phosphorylated neurofilament heavy chain; AUC - area under the curve; CI - 95% confidence interval.

4.4.5 Correlation with clinical parameters

Disease progression rate

In first visit samples from patients with ALS and PLS, linear regression models of $\log_{10}\text{PR}$ were constructed. Higher levels of CHIT1 and CHI3L2 were associated with higher PR in both simple linear regression and multiple linear regression models controlling for age of onset, site of onset and gender (CHIT1 simple regression Pearson's $r = 0.49$, $p < 0.001$, multiple regression $p = 0.049$; CHI3L2 $r = 0.42$, $p < 0.001$, multiple regression $p = 0.048$; **Figure 4.4A** and **C**. Multiple regression models are detailed in **Table 4.3**). CHI3L1 levels did not correlate with the rate of disease progression in simple or multiple regression models (simple regression $r = 0.13$, $p = 0.241$; multiple regression $p = 0.176$; **Figure 4.4B** and **Table 4.3**).

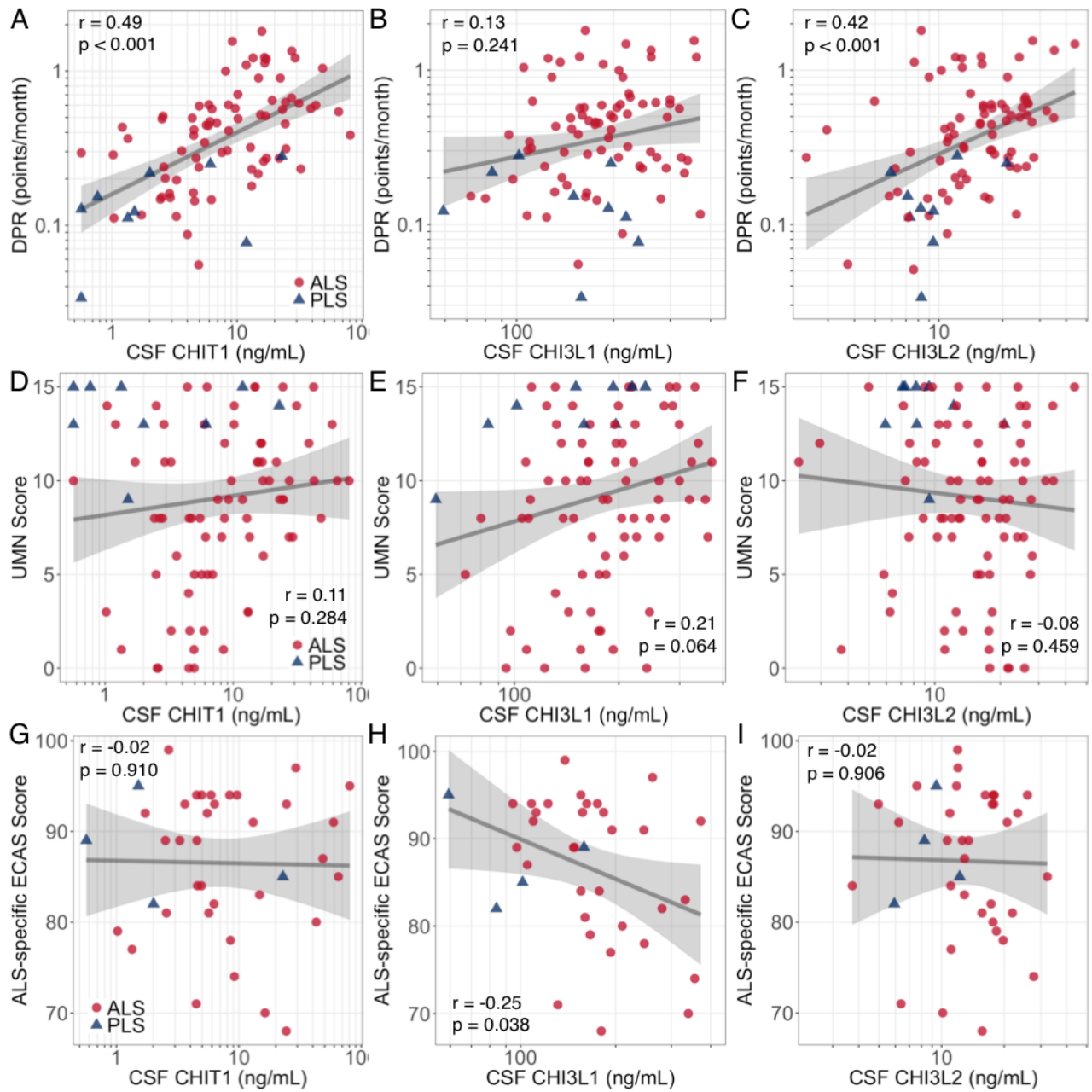


Figure 4.4: Correlation of CHIT1 (A, D, G), CHI3L1 (B, E, H) and CHI3L2 (C, F, I) with PR (A-C), UMN score (D-F) and cognitive function (G-I) in patients with ALS and PLS. Line indicates simple linear regression fit $\pm 95\%$ CI (grey ribbon). ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis; CHIT1 - Chitotriosidase-1; CHI3L1 - Chitinase-3-like protein 1; CHI3L2 - Chitinase-3-like protein 2; PR - rate of disease progression; UMN - upper motor neuron; ECAS - Edinburgh Cognitive and Behavioural ALS Screen.

	Chitotriosidase 1			Chitinase-3-like protein 1			Chitinase-3-like protein 2		
	Gradient	CI	<i>p</i>	Gradient	CI	<i>p</i>	Gradient	CI	<i>p</i>
$\log_{10}$ chitinase protein	0.264	0.002-0.527	0.049	0.621	-0.234-1.526	0.176	0.620	0.005-1.234	0.048
Sex (male)	0.041	-0.234-0.326	0.774	-0.068	-0.400-0.264	0.684	-0.114	-0.440-0.211	0.486
Site of onset (limb)	0.133	-0.178-0.444	0.398	0.200	-0.170-0.571	0.285	0.147	-0.204-0.500	0.406
Age of onset (years)	0.014	0.005-0.023	0.004	0.015	0.004-0.026	0.008	0.014	0.003-0.025	0.011
Adjusted R^2	0.10		0.015	0.07		0.041	0.10		0.014

Table 4.3: Multiple regression models of $\log_{10}$ PR of patients with ALS and PLS by chitinase protein level, controlling for the effects of gender, site of onset and age of onset. CI - 95% confidence interval; PR - disease progression rate; ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis.

Upper motor neuron signs

Similar regression analysis was performed to correlate $\log_{10}$ transformed chitinase levels with the burden of UMN signs (see **Section 2.3**). In contrast to the correlations with PR, higher CHI3L1 was weakly associated with higher UMN score in multiple linear regression models controlling for the degree of disability as measured by the ALSFRS-R and the disease duration, though not in simple linear regression (simple regression Pearson's $r = 0.21$, $p = 0.064$; multiple regression $p = 0.048$, **Table 4.4** and **Figure 4.4E**); there was no association between CHIT1 or CHI3L2 and UMN score in either simple or multiple regression models (CHIT1 $r = 0.11$, $p = 0.284$, multiple regression $p = 0.284$; CHI3L2 $r = -0.08$, $p = 0.459$, multiple regression $p = 0.442$; **Table 4.4** and **Figure 4.4D, F**).

	Chitotriosidase 1			Chitinase-3-like protein 1			Chitinase-3-like protein 2		
	Gradient	CI	<i>p</i>	Gradient	CI	<i>p</i>	Gradient	CI	<i>p</i>
$\log_{10}$ chitinase protein	2.027	-0.082-4.235	0.059	5.694	0.055-11.33	0.048	-1.61	-5.739-2.528	0.442
ALSFRS-R	-0.061	-0.264-0.142	0.554	-0.097	-0.282-0.088	0.300	-0.152	-0.338-0.035	0.110
Disease duration	0.260	0.013-0.507	0.039	0.159	-0.061-0.378	0.154	0.086	-0.147-0.319	0.465
Adjusted R^2	0.05		0.061	0.07		0.041	0.03		0.146

Table 4.4: Multiple regression models of UMN score of patients with ALS and PLS by chitinase protein level, controlling for the effects of gender, site of onset and age of onset. CI - 95% confidence interval; UMN - upper motor neuron; ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis.

Cognitive function

Regression analysis was also performed on a subgroup of patients ($n = 39$) with ALS and PLS who had undergone testing using the ECAS (see **Section 2.3** for further details). Higher levels of CHI3L1 were found to be weakly associated with a greater degree of cognitive dysfunction as measured by the ALS-specific ECAS score in both simple linear regression and multiple linear regression models controlling for the effect of disease duration and the degree of disability as measured by the ALSFRS-R (Pearson's $r = -0.25$, $p = 0.038$; multiple regression $p = 0.044$; **Table 4.5** and **Figure 4.4H**). Neither CHIT1 nor CHI3L2 levels were associated with ALS-specific ECAS score in simple or multiple regression models (CHIT1 Pearson's $r = -0.02$, $p = 0.910$, multiple regression $p = 0.811$; CHI3L2 Pearson's $r = -0.02$, $p = 0.906$, multiple regression $p = 0.967$; **Table 4.5** and **Figure 4.4G, I**).

	Chitotriosidase 1			Chitinase-3-like protein 1			Chitinase-3-like protein 2		
	Gradient	CI	p	Gradient	CI	p	Gradient	CI	p
$\log_{10}$ chitinase protein	-0.737	-6.949-5.474	0.811	-15.366	-30.311 to -0.421	0.044	0.290	-14.044-14.624	0.967
ALSFRS-R	0.163	-0.409-0.736	0.565	0.128	-0.423-0.678	0.640	0.232	-0.343-0.806	0.419
Disease duration	0.021	-0.611-0.653	0.947	-0.012	-0.558-0.533	0.963	0.063	-0.525-0.652	0.828
Adjusted R^2		-0.08	0.934		0.05	0.050		-0.07	0.869

Table 4.5: Multiple regression models of cognitive function, as measured by the ALS-specific ECAS score, of patients with ALS and PLS by chitinase protein level, controlling for the effects of gender, site of onset and age of onset. CI - 95% confidence interval; UMN - upper motor neuron; ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis.

4.4.6 Correlation of biochemical parameters

Chitotriosidase activity In order to test whether the CHIT1 detected in CSF is active, CSF chitotriosidase activity was measured with a fluorimetric activity assay. Chitotriosidase activity was highly correlated with CHIT1 levels (data from all visits; Pearson's $r = 0.93$, $p < 0.001$; **Figure 4.5A**), suggesting that elevated levels of CHIT1 do not constitute inactive enzyme.

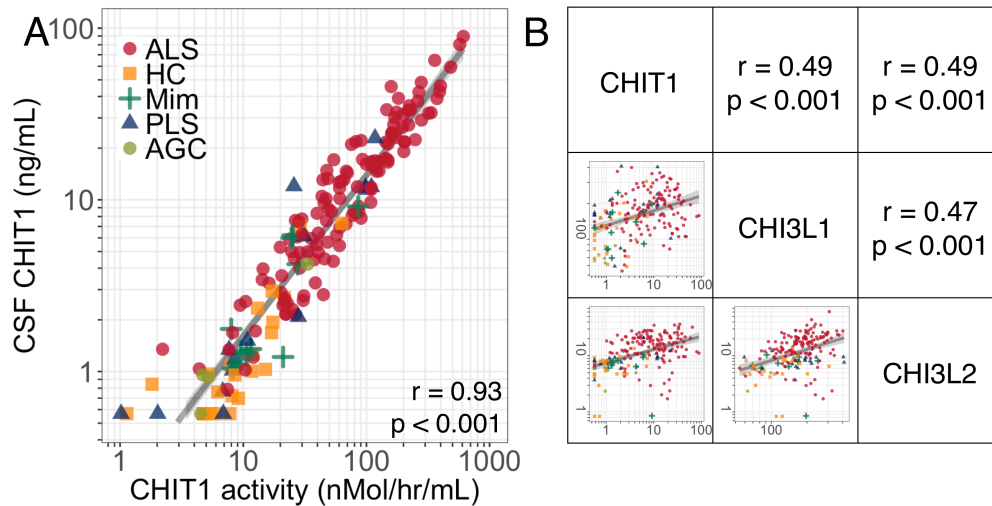


Figure 4.5: Correlation of CSF chitinase levels and activity. **A** Correlation of chitotriosidase activity and CHIT1 levels **B** Correlation between chitinase proteins. Line indicates simple linear regression fit $\pm$ 95% CI (grey ribbon). Pearson correlation coefficient indicated. ALS - amyotrophic lateral sclerosis; HC - healthy control; PLS - primary lateral sclerosis; Mim - mimic conditions; AGC - asymptomatic gene carrier; CHIT1 - Chitotriosidase-1; CHI3L1 - Chitinase-3-like protein 1; CHI3L2 - Chitinase-3-like protein 2.

Chitinase proteins Levels of the three chitinase proteins were moderately correlated (data from all visits; CHIT1-CHI3L1 Pearson's $r = 0.49$, $p < 0.001$; CHIT1-CHI3L2 Pearson's $r = 0.49$, $p < 0.001$; CHI3L1-CHI3L2 Pearson's $r = 0.47$, $p < 0.001$; **Figure 4.5B**).

pNFH Levels of the chitinase proteins were moderately correlated with the axonal degeneration marker pNFH, with CHIT1 most closely correlated, followed by CHI3L2 and CHI3L1 (CHIT1 Pearson's $r = 0.70$, $p < 0.001$; CHI3L1 $r = 0.58$, $p < 0.001$; CHI3L2 $r = 0.61$, $p < 0.001$).

CRP A weak correlation between the levels of the acute phase protein CRP and only CHI3L1 was observed (data for all visits; CHIT1 Pearson's $r = -0.03$, $p = 0.647$; CHI3L1 $r = 0.24$, $p < 0.001$; CHI3L2 $r = 0.05$, $p = 0.523$; **Figure 4.7**).

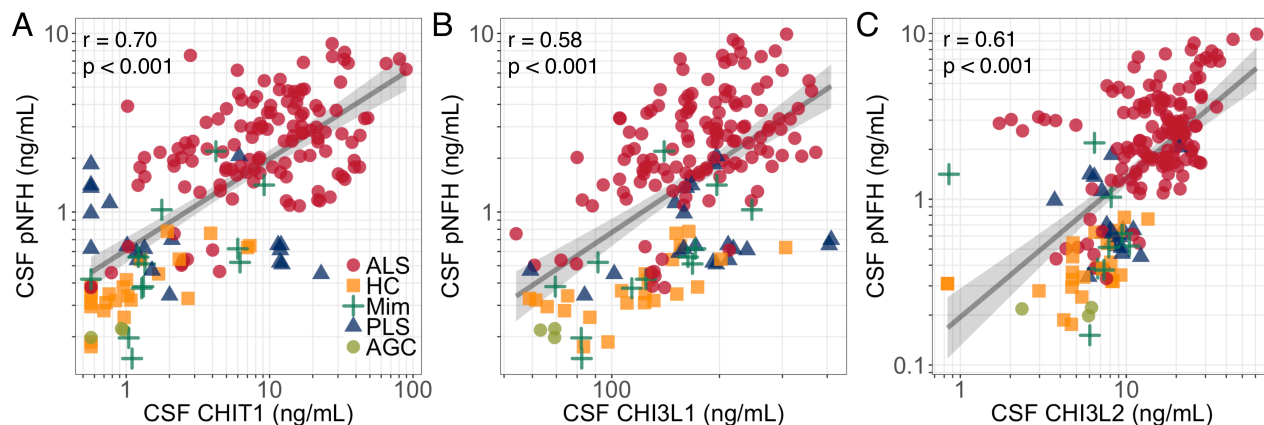


Figure 4.6: Correlation of CSF pNFH levels with **A** CHIT1, **B** CHI3L1 and **C** CHI3L2. Line indicates simple linear regression fit $\pm$ 95% CI (grey ribbon). Pearson correlation coefficient indicated. ALS - amyotrophic lateral sclerosis; HC - healthy control; PLS - primary lateral sclerosis; Mim - mimic conditions; AGC - asymptomatic gene carrier; CHIT1 - Chitotriosidase-1; CHI3L1 - Chitinase-3-like protein 1; CHI3L2 - Chitinase-3-like protein 2.

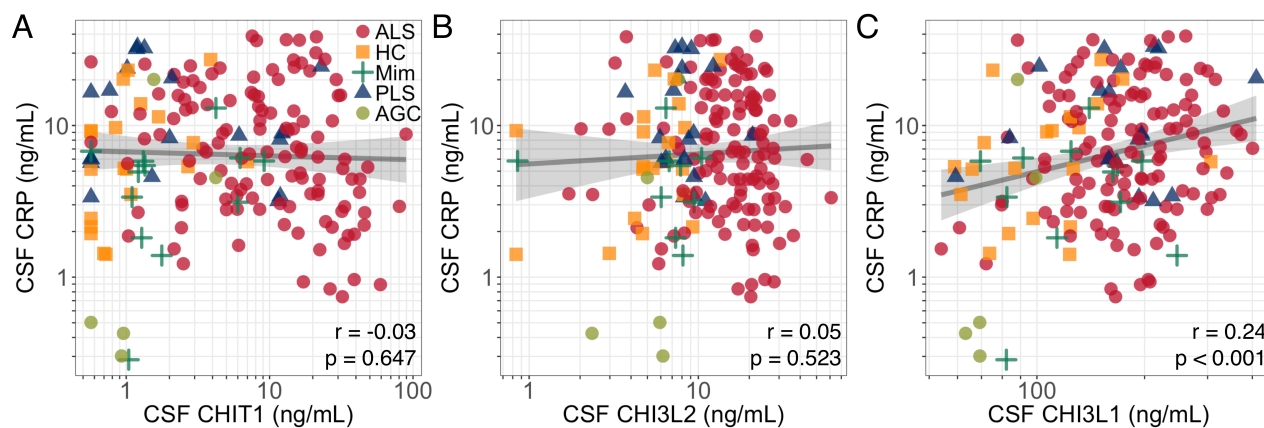


Figure 4.7: Correlation of CSF CRP levels with **A** CHIT1, **B** CHI3L1 and **C** CHI3L2. Line indicates simple linear regression fit $\pm$ 95% CI (grey ribbon). Pearson correlation coefficient indicated. ALS - amyotrophic lateral sclerosis; HC - healthy control; PLS - primary lateral sclerosis; Mim - mimic conditions; AGC - asymptomatic gene carrier; CHIT1 - Chitotriosidase-1; CHI3L1 - Chitinase-3-like protein 1; CHI3L2 - Chitinase-3-like protein 2.

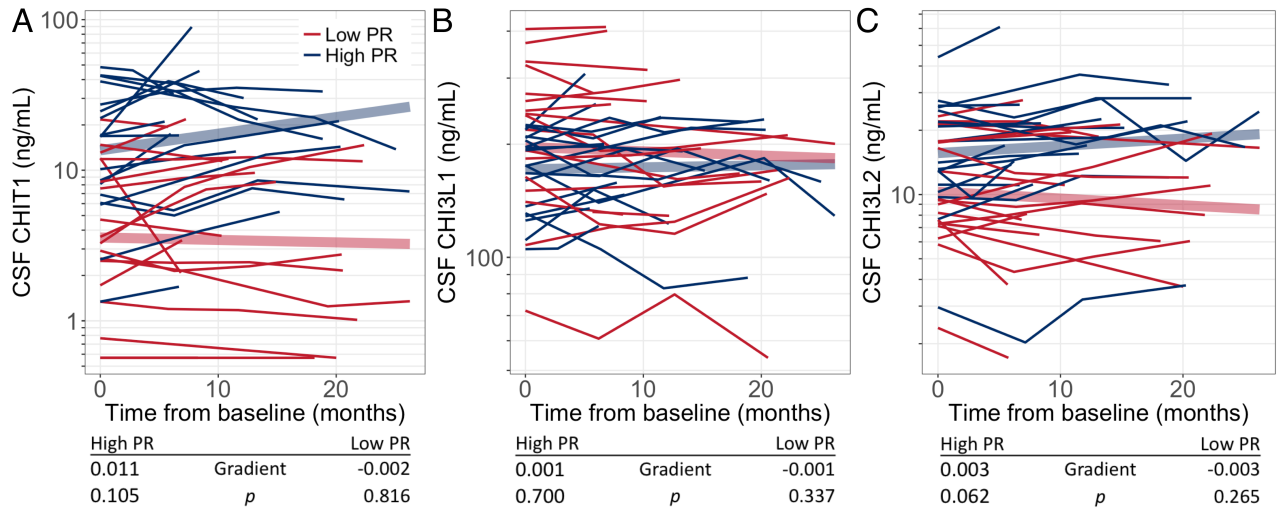


Figure 4.8: Longitudinal stability of CSF **A** CHIT1, **B** CHI3L1 and **C** CHIT1 in ALS and PLS patients, stratified by PR at first visit. Thick lines indicate group trend. ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis; CHIT1 - Chitotriosidase-1; CHI3L1 - Chitinase-3-like protein 1; CHI3L2 - Chitinase-3-like protein 2; PR - rate of disease progression.

4.4.7 Longitudinal analysis

Longitudinal CSF chitinase protein levels in ALS and PLS patients were fitted to a random intercept random slope linear mixed effects model, beginning at the baseline visit. The number of samples included in the longitudinal analysis is summarised in **Table 4.1**. Subjects were stratified into high- or low-PR categories dependent upon whether they were above or below (or equal to) the median PR at the baseline visit. There was no significant rise in the levels of any of the chitinase proteins over time whether stratified by PR or not (**Figure 4.8**).

4.4.8 Survival analysis

In univariate analysis, no significant association of any of the chitinase proteins and survival of MND patients was observed (CHIT1 $p = 0.075$; CHI3L1 $p = 0.922$; CHI3L2 $p = 0.239$; **Figure 4.9**). A Cox proportional hazards model was created, incorporating several factors known to be associated with shortened survival in ALS

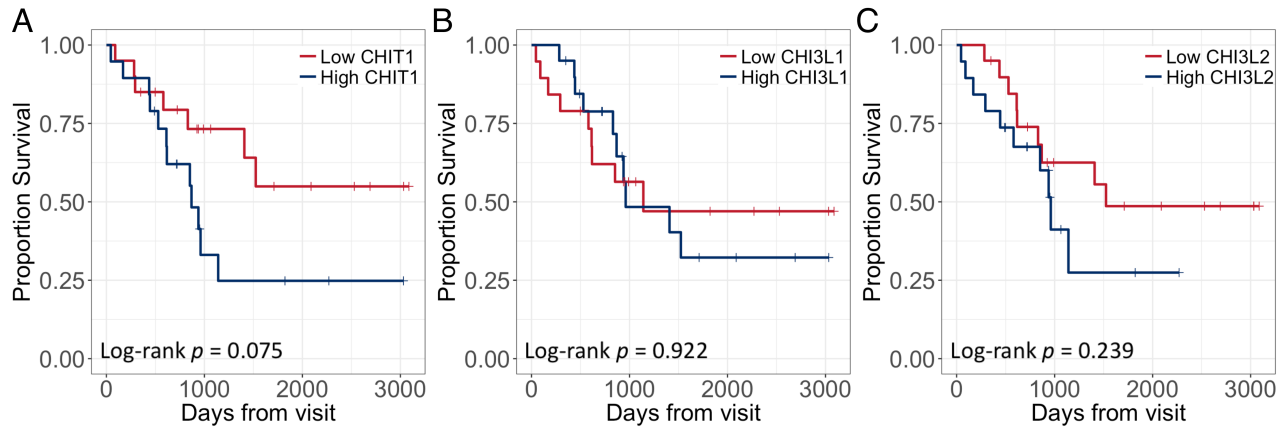


Figure 4.9: Kaplan-Meier curves for CSF **A** CHIT1, **B** CHI3L1 and **C** CHI3L2 in ALS and PLS patients, stratified by initial level at first visit. ALS - amyotrophic lateral sclerosis; PLS - primary lateral sclerosis; CHIT1 - Chitotriosidase-1; CHI3L1 - Chitinase-3-like protein 1; CHI3L2 - Chitinase-3-like protein 2.

(age of symptom onset, site of onset and gender) with chitinase proteins. Several additional factors, known to be associated with survival including respiratory function and *C9orf72* status were not included since the data was not available (Westeneng et al. 2018). This model identified a significant association of elevated CHIT1 (as well as older age of onset) with increased hazard of death (CHIT1 HR = 11.684, 95% CI 2.620-52.101, $p = 0.001$). On inclusion of pNFH in the model, there remained significant associations of CHIT1 and age with survival (CHIT1 HR = 6.831, 95% CI 1.135-105.765, $p = 0.021$; **Table 4.6**).

	HR	95% CI	p	HR	95% CI	p
Gender (male)	0.643	0.230-1.794	0.398	0.526	0.175-1.574	0.251
Onset site (spinal)	1.824	0.421-7.895	0.421	2.753	0.430-17.607	0.285
Age at onset (years)	1.105	1.023-1.192	0.011	1.090	1.011-1.176	0.025
Chitotriosidase 1	11.684	2.620-52.101	0.001	6.831	1.135-34.687	0.021
Chitinase-3-like protein 1	1.130	0.031-40.980	0.947	1.929	0.035-105.765	0.748
Chitinase-3-like protein 2	0.058	0.001-2.228	0.126	0.012	0.000-1.310	0.065
pNFH	-	-	-	1.338	0.917-1.011	0.118

Table 4.6: Cox proportional hazards survival modelling of CSF chitinase protein levels, with and without pNFH incorporating the effects of gender, site of onset and age of onset in ALS patients. HR - hazard ratio; CI - confidence interval; pNFH - phosphorylated neurofilament heavy chain.

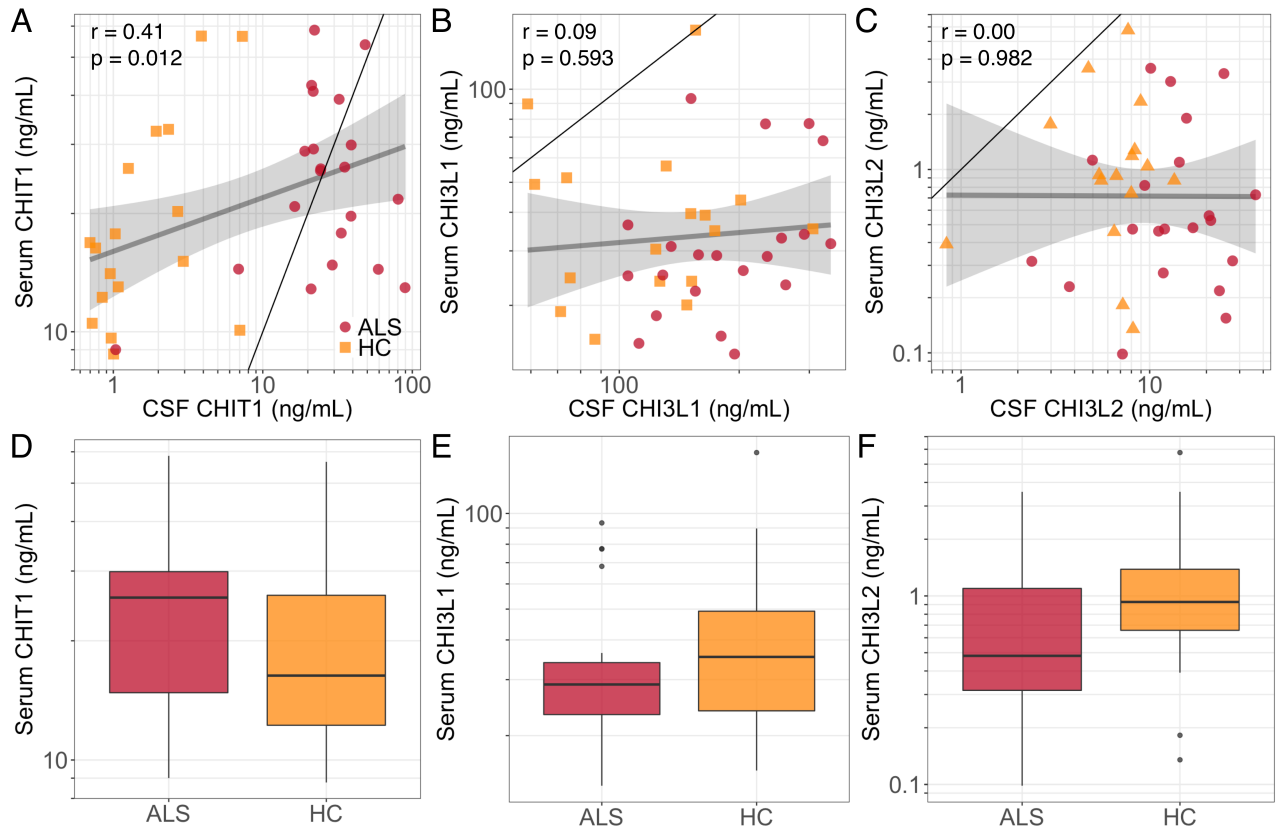


Figure 4.10: Levels of chitinase proteins in serum in a subset of ALS patients and healthy control subjects. **A-C** correlation of chitinase proteins in CSF and serum. Grey line indicates linear regression fit (grey ribbon 95% CI). Pearson's r indicated. Black line indicates CSF=serum. **A, D** CHIT1, **B, E** CHI3L1 and **C, F** CHIT1. ALS - amyotrophic lateral sclerosis; HC - healthy control; CSF - cerebrospinal fluid; CHIT1 - Chitotriosidase-1; CHI3L1 - Chitinase-3-like protein 1; CHI3L2 - Chitinase-3-like protein 2.

4.4.9 Serum chitinase proteins

Serum chitinase protein levels were measured in a cross-sectional subset of 20 ALS patients and 20 healthy control subjects. A weak correlation between CHIT1 levels in CSF and serum was observed, though there was no correlation between CSF and serum levels of CHI3L1 or CHI3L2 (CHIT1 Pearson's $r = 0.41$, $p = 0.012$; CHI3L1 $r = 0.09$, $p = 0.593$; CHI3L2 $r = 0.00$, $p = 0.982$; **Figure 4.10A-C**). No significant difference between serum levels of ALS patients and healthy controls was observed for any of the three chitinase proteins (**Figure 4.10D-F**).

4.5 Discussion

Overall, results of the analysis of the ELISA data described in this chapter provide validation of the main findings described in the proteomic analysis of **Chapter 3**: CSF chitinase proteins are elevated in ALS, correlate with clinical parameters, measures of axonal degeneration and survival; but diagnostic accuracy and associations with PR and survival do not surpass those of neurofilaments.

Comparing the two analyses in more detail yields a number of small differences. Firstly, the elevation of CHIT1 in ALS compared to PLS identified in proteomic analysis was not recapitulated in the analysis of *independent* samples (though it was found when data from all samples were analysed). By the same token, the similarity of CHI3L1 levels in ALS and PLS was not confirmed in *independent* samples (though was present when all samples were analysed). The earlier results suggesting detectable biochemical differences between ALS and PLS CSF samples therefore seem doubtful and might reflect differences in the latency from symptom onset to sampling in the earlier proteomic analysis.

Secondly, survival analysis of the combined cohort demonstrated an association of elevated CHIT1 with shortened survival in multivariate, though not univariate, analysis. This association was preserved with the inclusion of pNFH, unlike the earlier proteomic analysis. Further work including additional survival-associations would be helpful to determine whether this represents a genuine association of elevated CHIT1 independent of pNFH; the implications of this, perhaps suggesting that microglial activity is therefore a predictor of survival independent of axonal degeneration, are potentially important for therapy.

Finally, longitudinal analysis, rather than showing a gradual increase in CHI3L1 as shown in proteomic analysis, showed no change over time for any of the chitinase proteins.

These slightly different results in comparison with **Chapter 3** might be attributable in part to differences in the clinical characteristics of subjects included in the analysis. Alternatively, the differences could relate to the analytical methods used, as suggested by the imperfect correlation of the measurement of the three proteins between proteomic and ELISA quantification.

There are several potential reasons for discrepancies in quantification between the two studies. In the proteomic analysis, equal quantities of peptide were injected into the mass spectrometer and data was normalised by log-transformed median protein abundance. The proteomic measurement is therefore susceptible to alterations in overall protein content due to, for example, blood-CSF barrier dysfunction, which occurs with ageing and in ALS (Leonardi et al. 1984; Garbuzova-Davis et al. 2014). The ELISA data described here provides absolute abundance measurement thereby (at least partially) circumventing this. Other reasons for the variation between techniques could relate to systematic inaccuracies in proteomic or ELISA quantification: CHI3L2, for example, was quantified in proteomic analysis using just one unique peptide, whilst the presence of isoforms or modified protein not captured or detected by the antibodies used in the ELISA, or nonspecific antibody binding might cause inaccuracy in the ELISA data (Lin et al. 2013).

In addition to validating the results of **Chapter 3**, this analysis provides insight into the usefulness of chitinase proteins as biomarkers.

Levels of CSF chitinases were elevated in ALS patients compared with asymptomatic gene carriers and did not differ from healthy controls, suggesting that the chitinase proteins are probably not elevated decades prior to the onset of symptoms. There are a number of caveats to this. The gene carrier group was small and not age-matched to the ALS patients or healthy controls. Although the penetrance of the *C9orf72* hexanucleotide repeat expansion is thought to be high (Majounie et al. 2012), the age of symptom onset for the asymptomatic gene carriers is not known, nor is the clinical phenotype following symptom onset (i.e. ALS, FTD or ALS-FTD). Future

work on longitudinal samples from gene carriers, particularly covering the transition to the symptomatic period, would shed further light on the biomarker potential of chitinase proteins and might contribute to our understanding of the role of glia in ALS pathogenesis.

Interestingly, the three proteins are more closely correlated with pNFH levels than they are with each other, despite their purported similar cellular origin. Unlike CHIT1 and CHI3L2, which correlate with the rate of disease progression, CHI3L1 instead correlates with the degree of cognitive dysfunction and (weakly) the extent of clinical upper motor neuron involvement. This provides some support for the notion that the cellular, or at least regional, origin of CHI3L1 might differ from that of CHIT1 and CHI3L2, for example being produced by astrocytes (as found in Alzheimer's disease (Querol-Vilaseca et al. 2017), or perhaps by brain, as opposed to spinal cord, glia. Further work examining *post mortem* specimens, building upon those of Steinacker et al. 2018, would help to elucidate the relationship of CSF chitinases with the pathological processes underlying ALS and clarify their usefulness as markers of glial activity.

Serum levels of CHI3L1 and CHI3L2 (and CHIT1 in ALS patients) were lower than CSF levels, suggesting that most of the CHI3L1 and CHI3L2 detected in CSF chitinases are produced intrathecally, supporting a link with the pathological processes of ALS. Whilst this may well also be the case for CHIT1 (at least in ALS patients), additional work to calculate the expected serum-to-CSF ratio of CHIT1 would be required for confirmation.

The poor correlation of CSF and serum chitinase levels and lack of a difference between ALS patients and controls argues against measurement of serum chitinase levels as an ALS biomarker. Further work on a larger scale using independent cohorts would be helpful to delineate this further.

Lower levels (and activity) of CHIT1 has been shown in subjects with a common du-

plication in the CHIT1 gene, found in around a third of Europeans (Boot et al. 1998; Oeckl et al. 2018). Unfortunately, CHIT1 genotype is not available in this cohort, but the effect of the polymorphism on CSF CHIT1 levels might favour measurement of CHI3L2 instead, which shows similar diagnostic performance and clinical correlations in our dataset and is without similar confounding polymorphism.

5. UFLC-derived CSF Extracellular Vesicle Origin and Proteome

5.1 Hypothesis and Aims of This Chapter

5.1.1 Hypothesis

The ultrafiltration-liquid chromatography (UFLC)-extracted CSF EV proteome differs in composition from the whole CSF proteome and is more reflective of CNS intracellular events relevant to neurodegenerative diseases.

5.1.2 Aims

1. Demonstrate that UFLC extracts structures with morphology and marker expression typical of EVs
2. Compare the performance of CSF EV extraction using UFLC and UC in terms of yield and purity
3. Use open access datasets to delineate the tissue, CNS structural and subcellular origin of the UFLC-extracted CSF EV proteome

5.2 Introduction

Despite interest in the potential role of EVs in neurodegeneration (outlined in **Section 1.3.2**) and the proximity of CSF to the cells affected by neurodegeneration, little work has explored *CSF* EVs as a source of neurodegenerative disease biomark-

ers; the studies that have mostly comprise candidate-driven analysis of proteins and miRNA, and no candidate marker has been successfully independently validated (Saman et al. 2012; Feneberg et al. 2014; Gui et al. 2015; Stuendl et al. 2016).

A major challenge in the study of CSF EVs is their relative paucity, numbering fewer than 10×10^9 per mL in adults (Tietje et al. 2014). Standard methods of EV extraction based on UC were originally developed for high volume applications such as extraction from cell culture medium and have poor EV yield and high inter- and intra-operator variability (Lobb et al. 2015; Thery et al. 2006).

Alternative approaches employing precipitation, though producing higher yield, preclude downstream analytical methods such as MS-based proteomic analysis due to the use of polyethylene glycol and co-precipitation of large quantities of contaminating protein (Lobb et al. 2015). Further developments such as the use of size exclusion chromatography have provided high yield from cell culture medium and plasma, but these methods have not been systematically compared in CSF EV extraction (Nordin et al. 2015; Lobb et al. 2015). UC remains the only method capable of separating exosomes from other types of EV, using different centrifugal forces or density gradient centrifugation (Kowal et al. 2016).

In addition, although some details regarding the cargo of CSF EVs is known - for example, that they contain proteins found in microglia, neurons, astrocytes and oligodendrocytes as well as proteins implicated in neurodegeneration - little is known about their CNS regional origin, the contribution of blood EVs to the CSF EV population and whether the CSF EV cargo relates to intracellular processes (Chiasserini et al. 2014).

The experiments described in this chapter use a technique based on size exclusion chromatography (UFLC), previously successfully employed for high-yield and high-purity EV extraction from cell culture conditioned medium (Nordin et al. 2015). This is compared with EV extraction using a standardised UC method in measures

of EV yield, purity and proteomic coverage using LC-MS/MS.

In order to better understand the utility of CSF EVs in studying neurodegenerative diseases, with a specific focus on biomarker development, the whole CSF and CSF EV-enriched proteomes are also interrogated to elucidate the tissue and CNS structural origin of the identified proteins.

5.3 Materials and Methods

5.3.1 Participants and Sampling

CSF samples were obtained from patients attending the neurosciences ward at the John Radcliffe Hospital in Oxford, UK (Oxford University Hospitals NHS Foundation Trust) for lumbar puncture as part of routine care. Subject recruitment, CSF sampling and sample processing were completed as described in **section 2.3**. Diagnoses included multiple sclerosis, idiopathic intracranial hypertension, narcolepsy, autoimmune encephalopathies, normal pressure hydrocephalus and peripheral neuropathy. Samples with visible bloodstaining or red cell count >200 cells per mm^3 were not used.

5.3.2 EV extraction

Samples were thawed on ice and combined. Pooled samples were divided into 8mL aliquots. All analyses comparing UC, UFLC and CSF was performed on paired pools in order to enable direct comparisons.

Pooled samples underwent centrifugation at $1200g$ for 10 minutes to remove any cell debris or large aggregates and were filtered through a $0.22\mu\text{m}$ Millex[®] 33mm polyethersulfone syringe-driven filter (Merck Millipore, UK) to remove larger mi-

crovesicles (MVs) and apoptotic bodies prior to ultrafiltration or UC.

5.3.3 Ultracentrifugation

UC was performed as described previously, with modifications (Thery et al. 2006). Following 0.22 μ m filtration, samples were centrifuged in an Optima™ Max ultracentrifuge with MLS-50 swinging bucket rotor (Beckman Coulter, UK; k-factor 71) at 120,000*g* for 120 minutes. The EV pellet was resuspended in 4mL phosphate buffered saline (PBS) and subjected to a further spin at 120,000*g* for 120 minutes. The pellet was then resuspended in 100 μ L PBS.

5.3.4 Ultrafiltration-liquid chromatography

Following 0.22 μ m filtration, samples were filtered using Amicon® Ultra-15 100kDa molecular weight cut-off (MWCO) centrifugal filters (Merck Millipore) at 3500*g* for 8 minutes, washed with 4mLs PBS and centrifuged at 3500*g* for a further 4 minutes to remove abundant proteins below 100kDa. The retentate volume was adjusted to 800 μ L with PBS prior to injection into a size exclusion column packed with sepharose 4 fastflow (mean particle size 90 μ m, exclusion limit M_r 3×10^7), 24mL volume and eluted with PBS at a rate of 0.5 mL/min to a total of 40 mLs using an ÄKTA™ pure chromatography system (GE Healthcare Life Sciences, UK). 2mL fractions were collected from 6 to 40 mLs elution volume. EV-containing fractions (fractions 2-3) were concentrated for further analysis using Amicon® 10kDa MWCO 4mL centrifugal filters.

5.3.5 EV characterisation

EV size distribution and concentration was ascertained by nanoparticle tracking analysis (NTA) using a NanoSight NS500 (Malvern Panalytical, UK) and NTA 2.3 software. Where necessary, samples were diluted in PBS to achieve a concentration of 2×10^8 - 2×10^9 particles per mL. The camera level was set to 14 and detection threshold 5. Three recordings of 30-60 seconds were obtained for each sample and estimations of size distribution were averaged across recordings.

5.3.6 Immunoblotting

Electrophoresis was performed using the XCell Surelock Mini-Cell. 19.5 μ L EV samples were mixed with 7.5 μ L NuPAGE[®] lithium dodecyl sulphate (LDS) Sample Buffer (4X; Invitrogen, UK) and 3 μ L NuPAGE[®] Reducing Agent (10X) and heated to 95°C for 8 minutes. Samples were subsequently loaded into 10-well 1.0mm NuPAGE[®] Novex[®] Bis-Tris 4-12% gels and run at 125V for 45-60 mins in MES sodium dodecyl sulphate (SDS) running buffer until the dye front reached the bottom of the gel. Proteins were then transferred to polyvinylidene fluoride membrane using the iBlot 2 gel transfer device (ThermoFisher Scientific, UK) for 7 minutes. Membranes were blocked for 60 minutes with gentle shaking at room temperature using 5% fat-free milk in Tris buffered saline with 0.1% Tween (TBS-T).

After blocking, membranes were incubated with primary antibody, anti-SDCBP, anti-PDCD6IP (Programmed cell death 6-interacting protein, also known as ALG-interacting protein, ALIX) or anti-CD9 (Abcam, UK) at 1:500 dilution overnight at 4°C with gentle shaking. Membranes were washed 3 times in TBS-T for 10 minutes each with gentle shaking prior to addition of secondary antibody, peroxidase-labelled anti-mouse (PDCD6IP) or anti-rabbit (CD9, SDCBP) at 1:5000 dilution for 60 minutes at room temperature with gentle shaking. Membranes were again washed

3 times in TBS-T for 10 minutes each with gentle shaking.

Membranes were then incubated with SuperSignalTM West Pico peroxidase chemiluminescent reagent (ThermoFisher Scientific, UK) for 5 minutes at room temperature before development using photographic film, using appropriate exposure times of up to 10 minutes.

5.3.7 Transmission electron microscopy

10 μ L of EV solution was applied to freshly glow-discharged carbon-coated 200 mesh copper grids for 2 minutes, blotted with filter paper and stained with 2% uranyl acetate for 10 seconds, then blotted and air dried. Grids were imaged in a FEI Tecnai 12 transmission electron microscope (ThermoFisher Scientific) at 120 kV using a Gatan OneView CMOS camera.

5.3.8 Sample preparation for proteomic analysis

Samples of whole CSF and concentrated UFLC and UC samples were reduced in 5mM DTT for 30 minutes at room temperature followed by alkylation with 20mM IAA for 30 minutes at room temperature. Samples were subsequently precipitated using chloroform-methanol precipitation (Wessel et al. 1984). Precipitated protein was resuspended in 50mM triethylammonium bicarbonate (TEAB) with vortexing and sonication for 2 minutes. Samples were digested overnight at 37°C at 300 rpm using 400ng of trypsin for UFLC and UC samples and a 1:50 trypsin:protein ratio for whole CSF samples. Peptide digests were acidified with 1% formic acid and desalted using SOLA SPE cartridges (ThermoFisher Scientific, UK) and dried by vacuum centrifugation. Peptides from UFLC and UC samples were resuspended in 10 μ L of buffer A (2% acetonitrile, 0.1% formic acid in water) and whole CSF samples were resuspended to a final concentration of 500 ng/ μ L and kept at -20°C until

analysis. 1 μ L digested CSF peptide and 5 μ L extracted EV peptide were analysed by LC-MS/MS.

5.3.9 Liquid chromatography-tandem mass spectrometry

Peptides were analysed by nUHPLC LC-MS/MS using a Dionex Ultimate 3000 nanoUPLC, (ThermoFisher Scientific, Germany) coupled to an Orbitrap Fusion LumosTM TribridTM mass spectrometer (Thermo Scientific, Germany). Samples of peptide digest were loaded onto an EASY-Spray column (75 μ m $\times$ 500 mm, 2 μ m particle size, ThermoFisher Scientific, Germany) and eluted using a 60-minute gradient starting with 2% acetonitrile with 0.1% formic acid and 5% DMSO and increasing to 35% acetonitrile with 0.1% formic acid and 5% DMSO at a flow rate of 250nL/min. The data were acquired in data dependent mode with a resolution of 120,000 full-width half maximum at m/z 200 in the survey scan (375-1500 m/z) and with EASY-IC using the reagent ion source (202 m/z) for internal calibration. MS/MS spectra were acquired after precursor isolation in the quadrupole with an isolation window of 1.2 Th, dynamic precursor selection (top speed mode) with a fixed duty cycle time of 3 seconds and dynamic precursor exclusion of 60 seconds. Isolated precursor ions were fragmented by CID with a normalised Collision Energy of 35%. Parallelization was enabled and MS/MS spectra were acquired in the linear ion trap for up to 250 ms with an ion target of 4000 in rapid scan mode.

Raw MS data were analysed using Progenesis QI for Proteomics software v3.0 (Non-linear Dynamics). MS/MS spectra were searched against the UniProt Homo Sapiens Reference proteome (retrieved 15/11/2016) using Mascot v2.5.1 (Matrix Science) allowing for a precursor mass tolerance of 10 ppm and a fragment ion tolerance of 0.5 Da. Carbamidomethylation on Cysteines was defined as fixed modification and variable modifications included Deamidation on Asparagine and Glutamine, and Oxidation on Methionine. The peptide FDR was set at 1% and all peptides with

an ion score higher than 20 were imported into Progenesis QIP.

5.3.10 Normalisation and statistical analysis

Proteins with at least 2 unique peptides were included in the comparison of UFLC and UC proteomes. Protein abundance was normalized by $\log_2$ -transformation followed by centering on the median abundance of the 90% of proteins with the lowest variance across all runs and scaling by median absolute deviation. Missing values were imputed by random sampling from a normal distribution downshifted 1.8 SDs from the mean intensity of normalised values, with SD 0.25 times the SD of normalised values (Keilhauer et al. 2015).

Statistical analysis was performed as described in **Section 2.4**. Since the normalised log abundance values followed a normal distribution, comparisons of abundance were performed using a Welch's t -test with FDR correction using the Benjamini-Hochberg step-up procedure. GO enrichment analysis was performed in R using TopGO using the weight algorithm and a target-background approach. In comparisons between UFLC and UC, proteins increased in UFLC or UC samples with FDR-adjusted $p < 0.05$ as the foreground list and all proteins identified as the background list. Conversion of UniProt accession, HUGO Gene Nomenclature Committee (HGNC) gene symbol and EntrezGene ID were performed using the R package myGene.

Matrix effects due to the major differences in composition of whole CSF and purified EVs are likely to have an effect on peptide ionization and quantification. Therefore, comparison of the whole CSF and CSF EV-enriched proteomes was performed using Mascot identifications, whereby the list of proteins identified by Mascot with at least one unique peptide in at least two whole CSF or EV samples were compared; proteins identified in whole CSF samples were excluded from the EV list.

Analysis of predicted protein secretion by classical and non-classical secretory path-

ways was performed using SecretomeP and SignalP (Bendtsen et al. 2004), with a SecretomeP score of >0.6 taken to indicate likely non-classical secretion, where an N-terminal signal peptide was not predicted by SignalP.

GO and overrepresentation analyses comparing EV and CSF proteomes were performed using a target-background approach at the gene level, with proteins identified by Mascot in only EV or whole CSF as the foreground list. The background list comprised all proteins identified in CSF and EV samples combined.

Tissue enrichment analysis was performed using data from the Human Protein Atlas, which compares RNA expression across 33 different organs (Uhlen et al. 2015). Overrepresentation of cerebral cortex-and liver-enriched proteins (mRNA > 5 fold higher than in the next highest tissue) was sought in EV-only or CSF proteins, using all tissue-enriched proteins as the background list, with a hypergeometric test for overrepresentation.

Brain structure enrichment analysis was performed using data from the Allen Brain Atlas microarray survey of the human brain (Hawrylycz et al. 2012), which includes normalized data from a total of 3702 samples from 6 brains of differing ages and genders using over 62,000 probes with at least two probes for 93% of genes with EntrezGene IDs. Details of donor brains, brain structure ontology and normalization strategy are available in Hawrylycz et al. (2012) and via the Allen Brain Atlas data portal (<http://www.brain-map.org>). Tissue samples were grouped according to the Allen Brain Atlas ontology to cerebral cortex, subcortical grey structures (diencephalon, mesencephalon, metencephalon, myelencephalon) and white matter (telencephalic white matter, mesencephalic white matter, metencephalic white matter and myelencephalic white matter).

Gene expression was compared between structures using the probe showing the maximum variance across all samples for each gene. Proteins showing higher mean expression in each structure (fold change > 1.2) were then compared with the EV-

only or CSF-only protein list, abstracted to genes, using the background of all proteins identified in the whole CSF and CSF EV proteomic analysis. Overrepresentation was analysed using a hypergeometric test. Of the selected structures, only cerebral cortex, diencephalon, mesencephalon, metencephalon, myelencephalon, telencephalic white matter and choroid plexus had differentially expressed genes and were therefore used in the overrepresentation analysis.

Overrepresentation of EV and CSF proteins in myelin, frontal cortical and hippocampal neurons was performed, using proteomic datasets obtained from previously published studies. Overrepresentation of CSF and EV proteins was sought by comparing the whole CSF and EV-enriched proteomes with the proteome identified using mass spectrometry in the myelin and laser capture microdissected cortical and hippocampal neurons (Ishii et al. 2009; Drummond et al. 2015; Hondius et al. 2016), using a hypergeometric test, compared with a background of all proteins identified in the whole CSF and CSF EV proteomic analysis.

5.4 Results

5.4.1 Characterisation of UFLC and UC performance

EV isolation and yield

NTA of eluted UFLC samples identified the peak particle concentration at 8-12mL elution volume, corresponding to fractions 2-3 and the terminal portion of the column void volume (**Figure 5.1A**). Most protein eluted at 16-30mL volume, corresponding to fractions 5-11 (**Figure 5.1A**). NTA of the concentrated EV peak fractions demonstrated particles with a size distribution characteristic of a mixed population of EVs (**Figure 5.1B**). Immunoblotting demonstrated characteristic EV markers CD9, SDCBP and PDCD6IP (**Figure 5.1C**). Transmission electron

microscopy (TEM) of concentrated fractions 2-3 demonstrated particles with size ranging from 26-173 nm diameter and morphology consistent with EVs (**Figure 5.1D**). Notably EVs of diameter $> 220\text{nm}$ would be excluded by $0.22\mu\text{m}$ filtration prior to UC or UFLC.

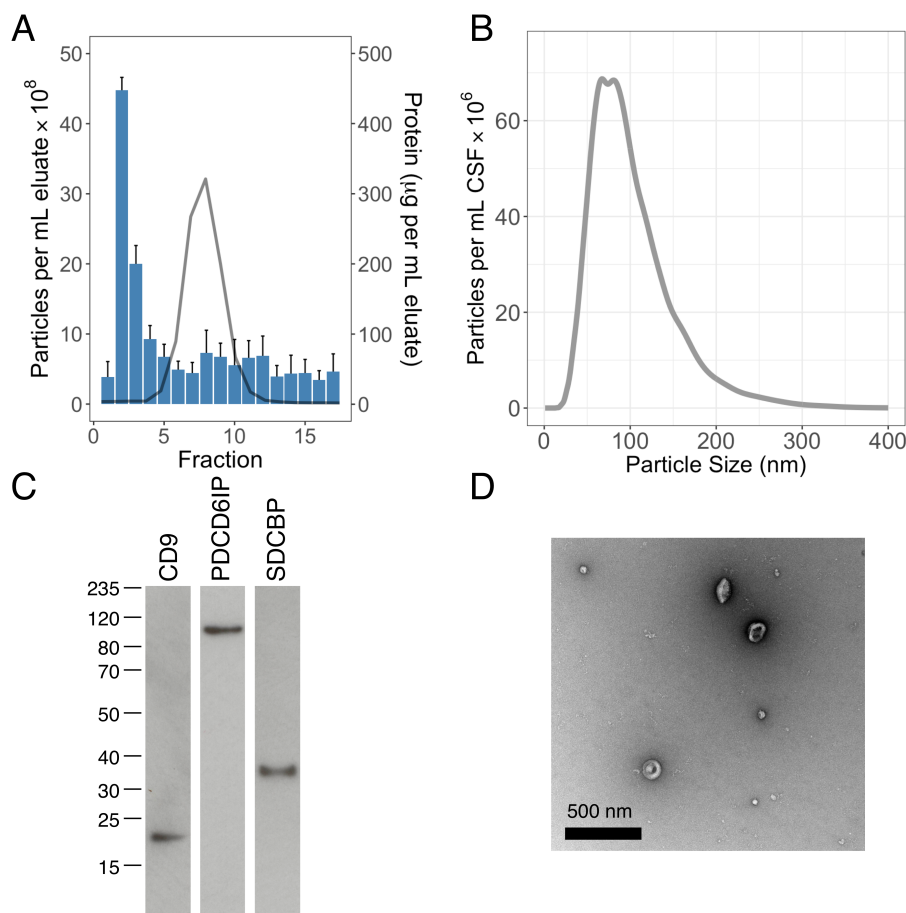


Figure 5.1: Extraction of EVs by UFLC. **A** Elution profile of UFLC. Blue bars represent NTA data of EV concentration per mL eluate (mean $\pm$ SD, $n=4$). Grey line indicates mean protein concentration in eluate (mean $\pm$ SD, $n=4$), demonstrating separation between EV peak (fractions 2-3) and protein peak (fractions 5-12) **B** Size distribution of extracted EVs following concentration measured by NTA, per mL of CSF starting volume (mean, $n=4$) **C** Western blot for EV markers PDCD6IP, CD9 and SDCBP in extracted EVs (fractions 2-3) **D** TEM of UFLC-extracted EV samples, demonstrating vesicles with a cup-shaped morphology. EV - extracellular vesicle; UFLC - ultrafiltration liquid chromatography; NTA - nanoparticle tracking analysis.

The total number of EVs recovered from 8mL CSF was over 7-fold greater using UFLC than UC (mean EVs recovered $6.32 \times 10^9 \pm \text{SD } 1.05 \times 10^9$ vs $0.85 \times 10^9 \pm 0.46 \times 10^9$, $p < 0.001$, $n=4$; **Figure 5.2A**) with comparable size distribution as measured by NTA (modal diameter UC mean = $96.25 \pm \text{SD } 19.69\text{nm}$, UFLC $76.25 \pm$

8.02nm, $p = 0.191$; combined median diameter UC = 111.88nm IQR 80.60-152.75nm, UFLC median 90.90nm, IQR 67.73-123.69nm; **Figure 5.2A**). The consistency of EV yield was also greater in UFLC compared to UC, with coefficients of variation 16.7% and 53.8% respectively. Characteristic EV markers including CD9, PDGFR α , SD-CBP, CD81, flotillin-1 and TSG101 were identified by LC-MS/MS consistently in both UFLC and UFLC isolated EVs. In keeping with the increased number of EVs demonstrated by NTA, relative abundance of all EV markers was higher in UFLC samples prior to normalisation ($p < 0.01$ for all markers).

Exclusion of contaminating protein

Estimation of protein amount using conventional assays was not possible prior to the proteomic analysis of EVs, due to the low protein concentration both in UFLC and UC samples. In order to examine the ability of the two techniques to exclude abundant CSF proteins, the ratios of the 10 most abundant CSF proteins to the most abundant EV marker CD9 were calculated. These were higher in UC for Albumin, Immunoglobulin Kappa constant, Cystatin C, Transthyretin, Alpha-1-antitrypsin and Transferrin, suggesting improved exclusion of abundant contaminating proteins using UFLC ($p < 0.05$, not adjusted for multiple comparisons **Figure 5.2C**). Interestingly, ratios of lipoprotein markers Apolipoprotein A1 and Apolipoprotein E were higher in UFLC, implying that the size-based extraction method is less effective at excluding lipoprotein particles such as low-density lipoprotein (LDL) and high-density lipoprotein (HDL) (**Figure 5.2C**). It should be noted, though, that some cell types have been reported to release Apolipoprotein E in association with exosomes (Niel et al. 2015), and other high-abundance CSF proteins might be released in EVs.

To extend this analysis, CD9 ratios of all proteins identified in whole CSF and EV samples (both UC and UFLC) were calculated (285 proteins). Overall, mean ratios were higher in UC than UFLC (UC mean $-1.44 \pm \text{SD } 1.32$, UFLC -1.75 ± 1.24 , $p =$

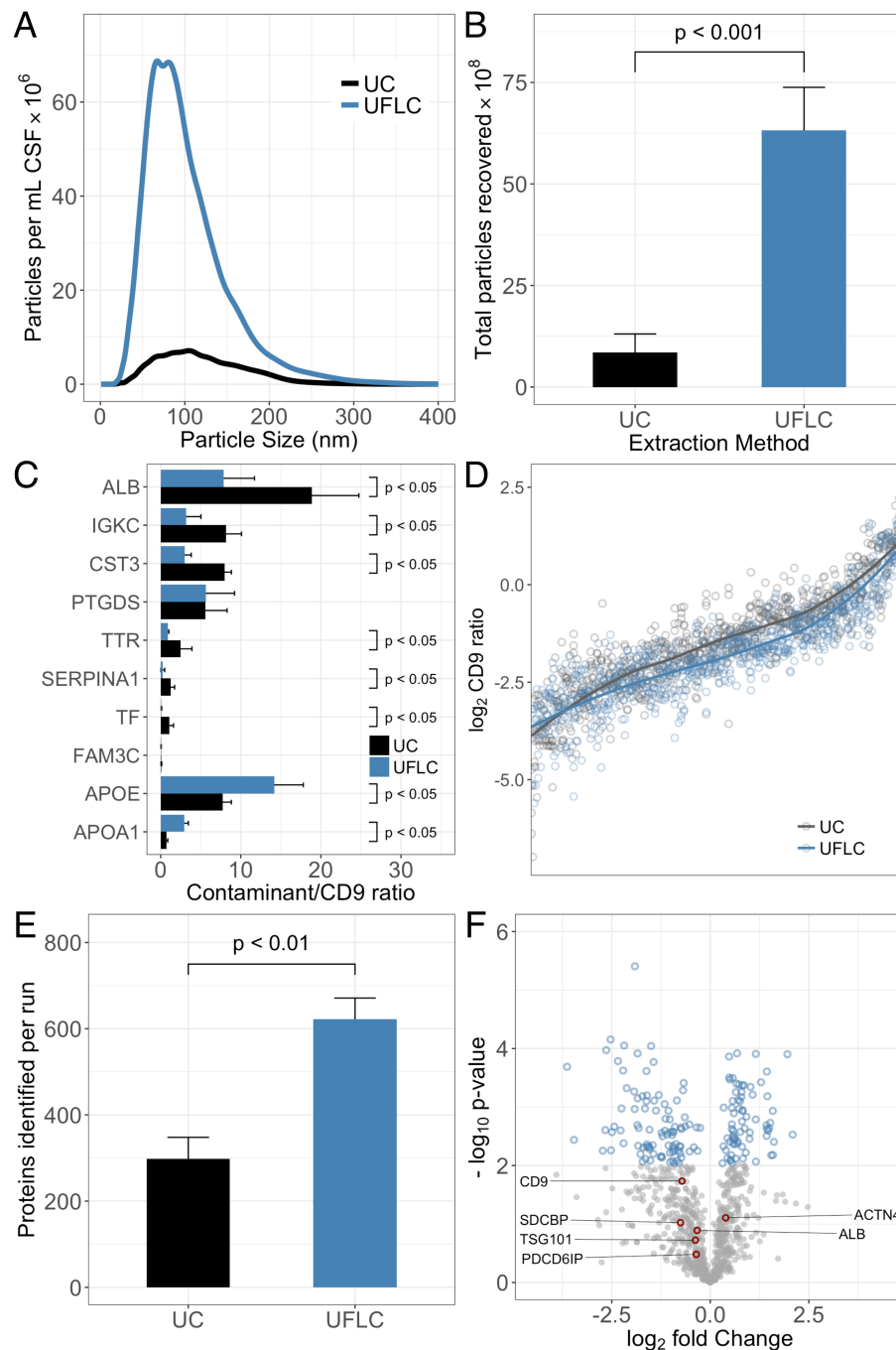


Figure 5.2: Comparison of UC and UFLC **A** Size distribution of UFLC-extracted EVs, following concentration of fractions 2-3 (upper blue line) and UC-extracted EVs (lower black line), per mL CSF (mean, n=4) **B** Total number of EVs recovered by UFLC and UC (mean $\pm$ SD, n=4) **C** Log ratio of top 10 most abundant CSF proteins to CD9 (n=3, mean $\pm$ SD) **D** CD9 ratio for all proteins identified in whole CSF and EV samples. Individual points indicate ratios for each UC or UFLC sample. Line indicates loess fit, proteins ordered by mean CD9 ratio. **E** Number of proteins identified per run of LC-MS/MS (n=3, mean $\pm$ SD) **F** Differences in protein abundance between UFLC and UC-extracted EV samples. Hollow blue points indicate proteins with FDR-adjusted $p < 0.05$. Hollow red points represent markers of different EV subtypes and albumin. ALB – Serum albumin; IGKC – Immunoglobulin kappa constant; CST3 – Cystatin-C; PTGDS – Prostaglandin-H2 D-isomerase; TTR – Transthyretin; SERPINA1 – Alpha-1-antitrypsin; TF – Transferrin; FAM3C – Protein FAM3C; APOE – apolipoprotein E; APOA1 – apolipoprotein A.

0.004), again implying improved exclusion of abundant proteins by UFLC (**Figure 5.2 D**).

UFLC and UC proteomes

A total of 1083 protein groups were identified with at least two unique peptides in UFLC and UC samples using the Mascot search engine. Label-free quantification of 668 proteins was possible across UFLC and UC samples (**Figure 5.2F**). The CV of label-free protein quantification in UFLC was lower than UC (median CV UFLC 16%, UC 21%). Significantly more Mascot identifications per run were achieved in UFLC samples than UC (704 ± 52 vs 340 ± 57 identifications, $p < 0.01$; **Figure 5.2E**). Following normalisation, 142 proteins were differentially abundant with FDR-adjusted $p < 0.05$ (**Figure 5.2F**). GO enrichment analysis of differentially abundant proteins compared with a background of all proteins identified in UC and UFLC samples indicated overrepresentation of terms (FDR-adjusted $p < 0.05$) “basal lamina” and “extracellular matrix structural component” in UFLC, and “intraciliary transport particle B” and “cilium morphogenesis” in UC.

Proteins that have been previously found to be enriched in EV populations isolated at different centrifugation speeds were examined for differential abundance in UC and UFLC samples (Kowal et al. 2016). There were no significant differences at the FDR-adjusted $p < 0.05$ threshold, though Alpha-actinin-4 (ACTN4), which is enriched in medium and large EVs pelleting at low speeds (2,000-10,000*g*), had higher mean abundance in UC samples, whilst syntenin-1, which is enriched in small EVs pelleting at higher speeds (100,000*g*), had higher mean abundance in UFLC samples (**Figure 5.2F**).

5.4.2 The CSF-EV proteome

Stringent criteria were applied for proteins in the EV list, with exclusion of proteins only identified in one sample and those detected also in CSF. The remaining 359 proteins were considered to constitute the EV-enriched proteome for subsequent analysis. All proteins identified in whole CSF (including those also found in EVs), identified in at least two CSF samples, were considered to constitute the CSF proteome, since proteins identified in both EV and CSF samples might represent contamination of EV samples with soluble whole CSF proteins.

Comparison of the EV-enriched proteome with databases of previously identified EV proteins, Exocarta and Vesiclepedia, demonstrated a high degree of overlap, with only 28 proteins identified in the EV-enriched proteome that were not included in Vesiclepedia or Exocarta (**Figure 5.3A**). Of the 100 most frequently identified EV proteins, including EV marker proteins, 45 were included in the EV-enriched proteome, compared with 22 identified in whole CSF (odds ratio (OR) 3.32, $p < 0.001$; **Figure 5.3B**).

GO enrichment analysis was performed, comparing EV-enriched proteome, CSF proteome and overlapping proteomes against the background of all proteins identified. Cellular component terms identified differences in the origin of CSF and EV-enriched proteins, with overrepresentation of extracellular vesicle terms, plasma membrane terms and cytoplasmic terms the EV-enriched proteome. CSF-enriched terms related to the extracellular space, complement and coagulation machinery ("membrane attack complex" and "platelet dense granule"; **Figure 5.4A-C**).

Biological process and molecular function terms in EVs related to extracellular vesicles, GTPase activity, and cell-cell adhesion, whilst those in CSF related to endopeptidases, extracellular matrix binding, complement, and humoral immunity. Notably, there was also some immune system-related terms also in the EV-enriched proteome

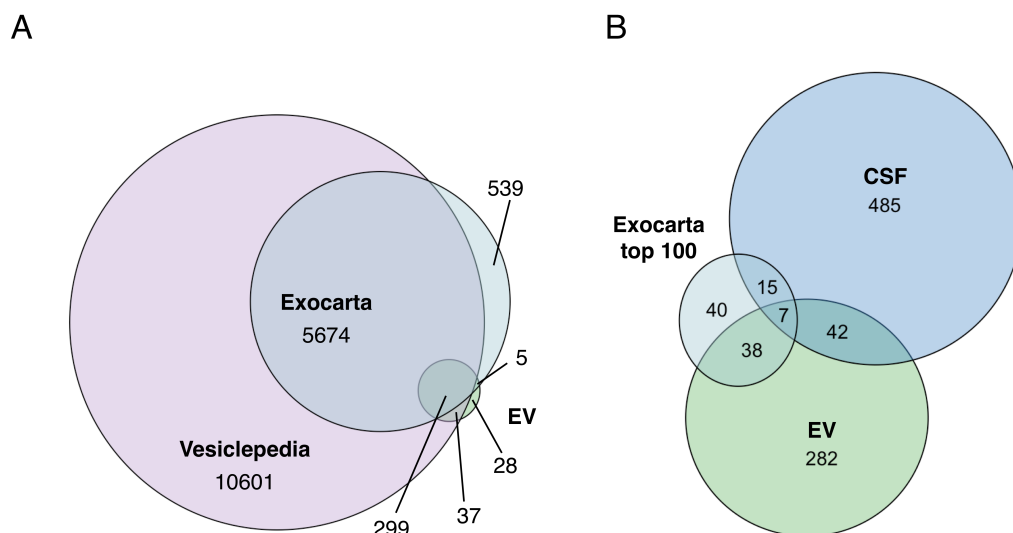


Figure 5.3: **A** Overlap of the EV-enriched proteome with EV protein databases Exocarta and Vesiclepedia. **B** overlap of CSF and EV proteomes with the top 100 most commonly identified proteins in exosome experiments according to Exocarta.

(“MHC class II complex binding”) and some intracellular terms in the whole CSF proteome (“cytoplasmic membrane-bound organelle lumen”). Interestingly, though not meeting the 5% FDR threshold, “myelin sheath” was enriched in the EV-enriched proteome, whilst “neuron recognition” and “synapse membrane” were enriched in the whole CSF proteome.

GO enrichment analysis of the overlapping proteome identified terms that also found in, or related to, whole CSF and EV-enriched GO terms. The enriched GO terms with FDR-adjusted $p < 0.05$ are given in **Figure 5.4**.

In agreement with GO enrichment analysis, the proportion of proteins with predicted intracellular origin (according to the Human Protein Atlas) was higher in the EV-enriched proteome compared to whole CSF (OR 2.63, $p < 0.001$) or the overlapping proteome (OR 2.49, $p < 0.001$; **Figure 5.5A**). In support of these findings, the proportion of predicted secretory proteins using SecretomeP and SignalP was higher in the CSF proteome compared with the EV-enriched proteome (OR 0.67, $p < 0.001$; **Figure 5.5B**); of proteins predicted to be secreted, those predicted to be non-classically secreted was much higher in the EV-enriched proteome compared

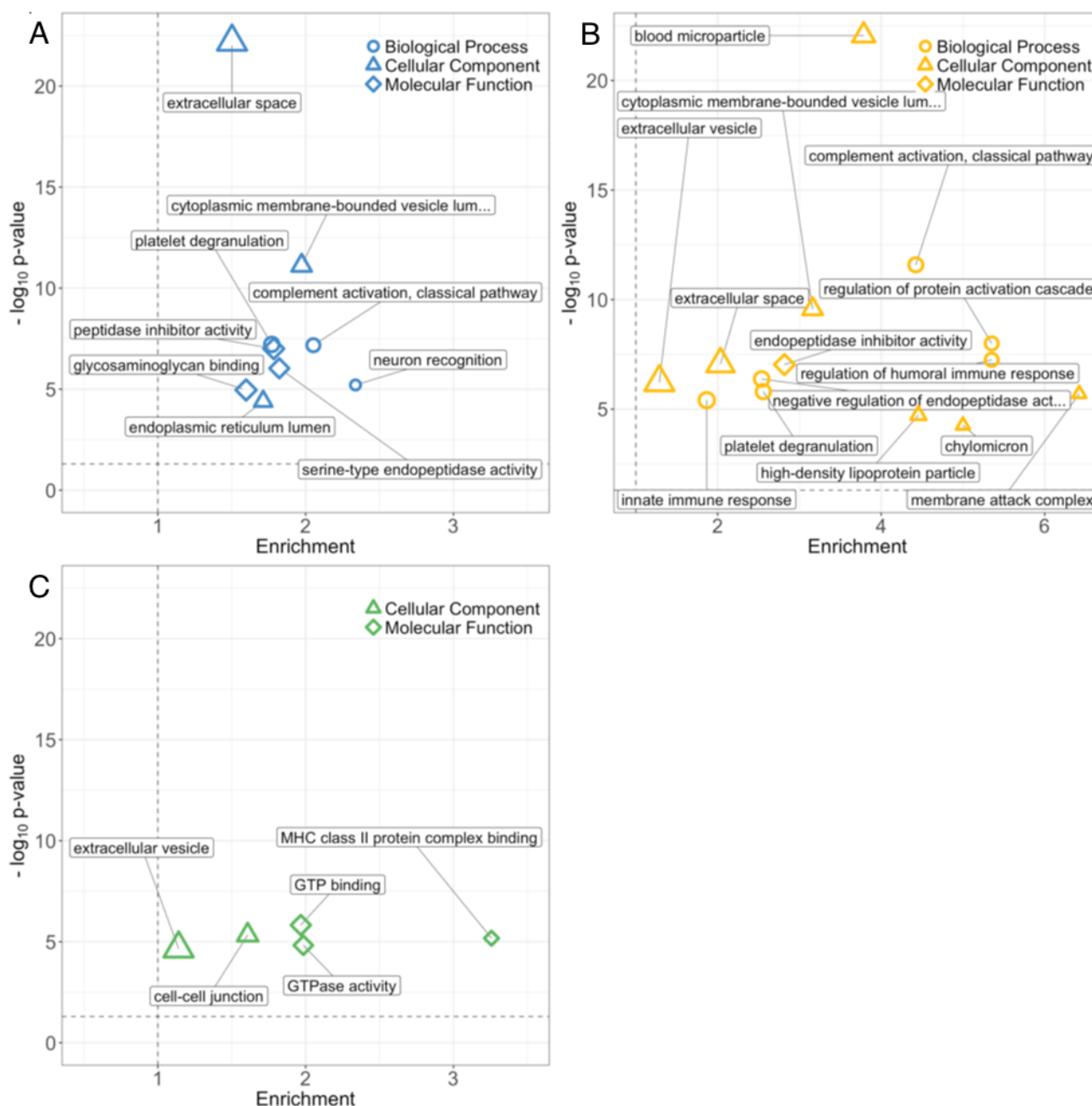


Figure 5.4: GO enrichment analysis of **A** whole CSF and **B** overlapping and **C** EV-enriched proteomes, using a background of all proteins identified in analysis. Points are proportional to the number of foreground genes annotated to individual terms. Vertical dashed line indicates no enrichment, horizontal line indicates $p = 0.05$. GO terms with FDR-adjusted $p < 0.05$ are shown.

with the CSF proteome (EV 8.19, $p < 0.001$; **Figure 5.5C**).

EV proteins were also less likely to be found in the plasma proteome compared to either whole CSF (OR 0.60, $p < 0.001$) or the overlapping proteome (OR 0.32, $p < 0.001$; **Figure 5.5D**), using data from the Human Protein Atlas.

Considering markers of specific organelles, the EV-enriched proteome contained the early endosome marker RAB5C (as well as eight other RAB proteins), multivesicular body markers CD81 and CD82 and the plasma membrane marker ATP1A1 but did not contain any of the endoplasmic reticulum (ER) markers calreticulin, calnexin and ER lumen protein-retaining receptor 1 (KDEL), or lysosomal proteins LAMP1, LAMP2A and LAMP2B.

Analysis of additional Human Protein Atlas data (which compares RNA expression in tissues from different organs to identify a tissue-enriched proteome) demonstrated overrepresentation of cerebral cortex-enriched proteins in both the EV and CSF proteomes (EV OR 3.18, $p < 0.001$; CSF OR 3.37, $p < 0.001$; **Figure 5.5E**). Twenty-two cerebral cortex-enriched proteins were identified in EV samples, 33 in whole CSF and 10 in the overlapping proteome. Overrepresentation analysis for liver-enriched proteins (a major contributor to the blood proteome through production of albumin, clotting factors and complement components; Uhlen et al. 2015), demonstrated a contribution of liver proteins to the overlapping (OR 44.83, $p < 0.001$), whole CSF (OR 19.47, $p < 0.001$) and, to a lesser extent, EV-enriched proteomes (OR 2.58, $p = 0.041$; **Figure 5.5C**).

To determine the relative contribution of different brain regions to the EV and CSF proteomes, analysis was performed using microarray data from the Allen Brain Atlas. Overrepresentation analysis of proteins with higher expression in brain structures was compared in the EV and whole CSF datasets, using all proteins identified in the proteomic analysis as a background list.

This analysis indicated overrepresentation of genes with higher expression in telen-

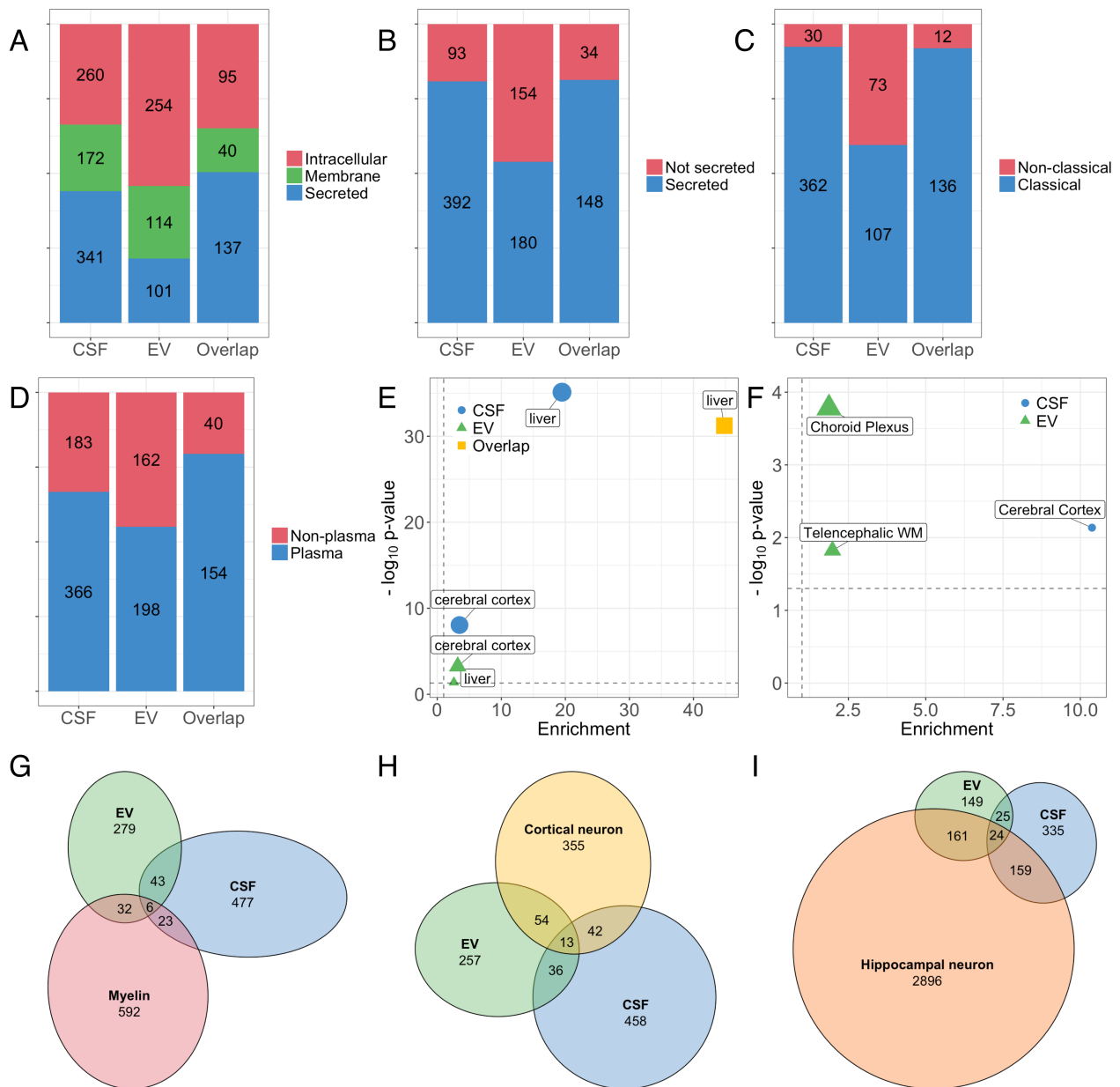


Figure 5.5: Origin of CSF EVs. **A** Predicted protein class identified proteins (note that some proteins are annotated to more than one class) **B** Predicted secreted and not secreted proteins using SecretomeP and SignalP **C** Predicted secreted and not secreted proteins using SecretomeP and SignalP **D** Proportion of proteins in the EV-enriched, whole CSF and overlapping proteins identified as plasma proteins. **E** Tissue and **F** brain structure enrichment analysis of whole CSF, overlapping and CSF EV-enriched proteomes. No brain structure enrichment was detected in the overlapping proteome. Point size proportional to the number of tissue enriched proteins per tissue. Vertical dashed line indicates no enrichment, horizontal line indicates $p = 0.05$. **G-I** Overlap of EV-enriched and CSF proteomes with proteomes of **G** myelin **H** laser capture microdissected frontal cortical neurons and **I** laser capture microdissected hippocampal neurons.

cephalic white matter in the EV-enriched proteome compared to other brain regions (OR 1.99, $p = 0.015$; **Figure 5.5F**). The genes overlapping between EV and telencephalic white matter expression predominantly comprised those expressed by oligodendrocytes or involved in myelination (including CA2, CD9, DNHD1, ENPP6, EPB41L2, HAPLN2, KLK6, PLEKHB1 and TMBIM1; Cahoy et al. 2008; Dugas et al. 2006; Kida et al. 2006; Parkkila et al. 2001; Terada et al. 2002; Xiao et al. 2016; Yang et al. 2015) as well as those expressed on axons (CA14; Parkkila et al. 2001), astrocytes (AQP1) and microglia (HLA-DRA Darmanis et al. 2015). Cerebral cortex proteins were overrepresented in the whole CSF proteome compared to the EV-enriched proteome (OR 10.38, $p = 0.007$). Choroid plexus proteins were also enriched in the EV-enriched proteome (OR 1.87, $p < 0.001$; **Figure 5.5F**).

There was no significant brain region enrichment identified in the overlapping proteome.

In order to provide additional validation for this finding, further overrepresentation analysis was performed using a dataset of 673 myelin proteins identified from human *post mortem* frontal lobe tissue in a previously published study (Ishii et al. 2009). This demonstrated significant enrichment in the EV but not whole CSF proteome, compared to the background of all proteins identified in the proteomic analysis, supporting the earlier enrichment analysis (EV OR 2.10, $p = 0.001$; CSF OR 0.65, $p = 0.977$, **Figure 5.5G**).

A similar analysis was performed in which overrepresentation of EV and whole CSF proteins was sought amongst 464 proteins identified in cerebral cortical neurons isolated using laser capture microdissection (LCM) in a previously published study (Drummond et al. 2015), in order to validate the finding of cerebral cortex enrichment in whole CSF. This did not support the earlier findings, instead also demonstrating enrichment of cerebral cortex proteins in the EV-enriched proteome (EV OR 1.38, $p = 0.006$; CSF OR 0.93, $p = 0.761$, **Figure 5.5H**).

Further analysis of an additional dataset comprising 3106 proteins identified in proteomic analysis of LCM-isolated hippocampal neurons from healthy controls and patients with Alzheimer’s disease also demonstrated overrepresentation of hippocampal proteins in the EV-enriched, but not whole CSF, proteome (EV OR 1.80, $p < 0.001$; CSF OR 0.57, $p = 0.999$; **Figure 5.5I**).

On examination of the dataset for current candidate neurodegenerative disease biomarkers, including amyloid beta A4 protein (APP), tau, S100 proteins, 14-3-3 proteins, neurogranin, presenilin 1 and 2, TDP-43, α -synuclein, Transforming growth factor beta (TGFB), chitinase proteins, major prion protein (PRNP) and neurofilament proteins (Lewczuk et al. 2018), only amyloid precursor protein, four S100 protein isoforms and four 14-3-3 protein isoforms were found in the EV-enriched list. CHI3L1 and TGFB were detected in the overlapping list, and PSEN1, CHI3L2, TREM2 and an additional isoform of APP were detected exclusively in whole CSF. In UC samples (including proteins overlapping with CSF), only APP, five 14-3-3 proteins and six S100 proteins were detected.

5.5 Discussion

The UFLC method described isolated particles with a size distribution and morphology similar to that extracted by UC and typical of EVs. Of note, UFLC vastly improved the particle and protein yield (around 7-fold) whilst also improving upon the purity of the EV proteome without systematic alterations in the identified proteins as analysed by GO enrichment. As a consequence of the improved yield, more proteins were identified in UFLC samples with less variable quantification, making this the method of choice for MS-based CSF EV biomarker discovery experiments. Although there appears to be co-isolation of apolipoprotein particles, this is not necessarily problematic for biomarker development in neurodegenerative diseases where alterations in lipoproteins have been described (Kim et al. 2009; Mariosa et al. 2017).

The main limitation of this method is the large volume of CSF required to obtain good quality LC-MS/MS data. Adaptations such as the use of ion exchange chromatography or employing smaller size exclusion column volumes may improve this. Other studies of the EV proteome employing different extraction and proteomic methods have used smaller volumes of CSF (5-6mLs), but due to the improved recovery and purity demonstrated here, UFLC remains preferable for biomarker discovery (Chiasserini et al. 2014; Welton et al. 2017).

Matrix effects, leading to differences in the behaviour of precursor ions from purified EV and whole CSF samples, render the comparison of the whole CSF and CSF EV proteomes difficult. For this reason, a semiquantitative approach based on the presence or absence of protein identifications was used to explore differences between the proteomes. Alternative methods such as the use of non-MS-based proteomics or isobaric labelling might avoid matrix effects and partially circumvent this issue, but are themselves associated with significant limitations (current aptamer-based proteomic technologies, for example, necessarily limit the proteins detectable to a prespecified list; Welton et al. 2017).

Analysis of the whole CSF and CSF EV proteomes using GO enrichment identified more intracellular terms in the EV proteome and more terms associated with the extracellular space and plasma in the whole CSF and overlapping proteomes. These findings support the enrichment analysis comparing predicted protein class and origin of the proteomes, suggesting that a greater proportion of the CSF EV proteome was of intracellular origin and that a greater proportion of the EV proteins that are secreted are secreted non-classically. This is in keeping with the hypothesis that the EV proteome contains biomarkers more reflective of intracellular events than the whole CSF proteome.

In addition, comparison with data from the Human Protein Atlas found a greater proportion of the whole CSF proteome to be found in plasma compared to the EV proteome. The overrepresentation of liver proteins in the whole CSF and overlapping

proteomes is in keeping with this and is likely to reflect the well-documented overspill of highly abundant plasma proteins into CSF (Guldbrandsen et al. 2014).

Overrepresentation analysis of proteins showing differences in regional expression was performed, showing enrichment of chroid plexus and white matter proteins in the CSF EV proteome and enrichment of cerebral cortical proteins in the whole CSF proteome. Although this analysis should be interpreted cautiously, since the regional expression patterns were based on RNA rather than protein data, the additional analysis of a myelin proteomic dataset provides support for the differential enrichment of myelin (i.e. white matter) proteins within EVs (including proteins such as 2'3'-cyclic-nucleotide-phosphodiesterase (CNP), previously identified in oligodendrocyte-derived EVs; Kramer-Albers et al. 2007). This data, though in need of validation, suggests that oligodendrocytes contribute proportionally more to the CSF EV proteome than their contribution to whole CSF, in contrast to other cell types. It should be noted, however, that samples from diverse conditions were included in the analysis, including CNS demyelination, in which increased oligodendroglial EV release might occur.

Interestingly, the enrichment of cortical proteins in whole CSF (and corresponding lack in the EV-enriched proteome) was not recapitulated in a proteomic dataset of neocortical and hippocampal neurons. There are several potential reasons for this discrepancy. For example, it may be a reflection of differences between the cellular transcriptome and proteome or relate to the preparation of samples for transcriptomic data (that they are tissue homogenates comprising multiple cell types and extracellular matrix rather than the relatively pure population of cells analysed in the proteomic dataset).

Several previously-established biomarkers of neurodegenerative diseases were identified in EV samples, though only S100 and 14-3-3 proteins were detected *exclusively* in EV samples, and a large proportion of current candidates were not detected by proteomic analysis at all, either due to low abundance or a lack of suitable peptides.

Notably, several proteins that are major constituents of aggregates in neurodegenerative diseases were not identified in this analysis (specifically TDP-43, tau and alpha-synuclein) but have previously been detected in CSF EVs in target-driven studies (Stuendl et al. 2016; Saman et al. 2012; Feneberg et al. 2014).

In conclusion, this chapter describes a method for extraction of CSF EVs that surpasses ultracentrifugation in terms of recovery and purity, thereby permitting increased proteomic depth. These data suggest that the EV-enriched proteome contains proteins expressed in multiple brain regions but in particular cerebral white matter and choroid plexus. Furthermore, this work demonstrates that analysis of CSF EVs increases proteomic depth of the intracellular compartment when compared to whole CSF and that EVs cargo represents components, processes and specific proteins relevant to neurodegenerative disease.

6. The Cerebrospinal Fluid Extracellular Vesicle Proteome in ALS

6.1 Hypothesis and aims

6.1.1 Hypothesis

Alterations in the CSF EV proteome of patients with ALS reflect cellular perturbations characteristic of the disease and can provide biomarkers of ALS

6.1.2 Aims

1. Extract and quantify the CSF EVs population from a cohort of ALS patients and relevant control subjects
2. Perform proteomic analysis of the extracted EVs:
 - i. Multivariate analysis using hierarchical clustering and PCA
 - ii. Compare ALS and control proteomes to identify candidate ALS biomarkers
 - iii. Compare *C9orf72* and non-*C9orf72*-associated ALS proteomes to explore dysregulated pathways

6.2 Introduction

As outlined in **Section 1.3.2** and **1.2.6**, EVs have been implicated in the pathogenesis of ALS, particularly as a mediator of intercellular spread of TDP-43 (Feiler et al. 2015; Thompson et al. 2016). Inhibition of exosome biogenesis leads to increases in insoluble TDP-43 in a TDP-43-mutant mouse model, and reduced EV secretion is observed in fibroblasts and iPSC-derived motor neurons from patients carrying the *C9orf72* hexanucleotide repeat expansion, (Iguchi et al. 2016; Aoki et al. 2017).

TDP-43 is also found at elevated levels within EV extracted from brain of patients with ALS (Iguchi et al. 2016); though a study using targeted MS to examine the quantity of TDP-43 in EVs extracted from the CSF of ALS and FTD patients using (and the only study of CSF EVs in ALS to date) did not recapitulate this finding in CSF, potentially due to the diverse origin of CSF EV and the ubiquity of TDP-43 expression (Feneberg et al. 2014).

Regardless of whether aberrant EV secretion or spread of pathological TDP-43 via EVs occurs in ALS, the experiments described in **Chapter 5** demonstrating increased insight into the intracellular compartment by studying the EV proteome suggest better access to pathways relevant to the pathogenesis of ALS than the study of whole CSF.

This chapter therefore describes the morphological characterisation and proteomic analysis of EVs extracted from individual CSF samples from a cohort of patients with ALS and relevant control subjects, using the optimised UFLC method described in **Chapter 5**.

	ALS	Mimic conditions	Healthy control	<i>p</i>
n, first visit	20	9	9	-
n, second visit	6	-	-	
n, third visit	4	-	-	
% male	75.0	100.0	55.6	0.082 [*]
Age at sampling, mean $\pm$ SD (years)	60.0 $\pm$ 10.9	54.5 $\pm$ 12.6	63.5 $\pm$ 15.1	0.245 ⁺
Age at onset, mean $\pm$ SD (years)	57.8 $\pm$ 10.7	-	-	-
% bulbar onset	25.0	-	-	-
Disease progression rate, median [IQR]	0.36 [0.18-0.58]	-	-	-

Table 6.1: Demographic data for the CSF EV cohort. ^{*} Fisher exact test. ⁺Kruskal-Wallis *H* test. ALS – amyotrophic lateral sclerosis; SD - standard deviation; IQR - interquartile range.

6.3 Materials and methods

6.3.1 Participants and sampling

This study used CSF from 20 patients with ALS (including 4 carrying the *C9orf72* hexanucleotide repeat expansion), 9 patients with mimic conditions (including multifocal motor neuropathy with conduction block, Kennedy’s disease, cervical spondylotic myeloradiculopathy and motor presentation of multiple sclerosis) and 9 healthy control subjects. Subject recruitment, CSF sampling, sample processing and clinical characterisation was completed as described in **Section 2.3**. Demographic and clinical characteristics are displayed in **Table 6.1**.

6.3.2 EV extraction

CSF samples were excluded if they had a red blood cell count $>200/\text{mm}^3$ or visible blood staining. EVs were extracted from 7.2mL CSF from each subject using UFLC as described in **Chapter 5**. Following UFLC, samples were concentrated using Amicon[®] Ultra-4 10kDa MWCO centrifugal filters (Merck Millipore) at 3500g to a retentate volume of $<400\mu\text{L}$. Sample order was randomised prior to EV extraction.

6.3.3 EV characterisation

EV size distribution and concentration was ascertained in fractions 2 and 3 by NTA using a NanoSight NS500 (Malvern Panalytical, UK) and NTA 2.3 software. Where necessary, samples were diluted in PBS to achieve a concentration of 2×10^8 - 2×10^9 particles per mL. The camera level was set to 14 and detection threshold 5. Three recordings of 30-60 seconds were obtained for each sample and estimations of size distribution were averaged across recordings.

6.3.4 Sample preparation for proteomic analysis

Samples were reduced in 5mM DTT for 30 minutes at room temperature followed by alkylation with 20mM IAA for 30 minutes at room temperature. Samples were subsequently precipitated using chloroform-methanol precipitation (Wessel et al. 1984). Precipitated protein was resuspended in 50mM TEAB with vortexing and sonication for 2 minutes. Samples were digested overnight at 37°C at 300 rpm using 400ng of trypsin. Peptide digests were acidified with 1% formic acid and desalted using SOLA SPE cartridges (ThermoFisher Scientific, UK) and dried by vacuum centrifugation. Peptide samples were resuspended in 10 μL of buffer A (2% acetonitrile, 0.1% formic acid in water). The entire sample of extracted EV peptide

was injected for analysis by LC-MS/MS. Sample order was randomised prior to MS analysis and a sample produced by pooling an equal volume of peptide from each sample was run after every 10 samples.

6.3.5 Liquid chromatography-tandem mass spectrometry

Peptides were analysed by nUHPLC LC-MS/MS using a Dionex Ultimate 3000 nanoUPLC, (ThermoFisher Scientific, Germany) coupled to an Orbitrap Fusion LumosTM TribridTM mass spectrometer (Thermo Scientific, Germany). Samples of peptide digest were loaded onto an EASY-Spray column (75 μm ×500 mm, 2 μm particle size, ThermoFisher Scientific, Germany) and eluted using a 60-minute gradient starting with 2% acetonitrile with 0.1% formic acid and 5% DMSO and increasing to 35% acetonitrile with 0.1% formic acid and 5% DMSO at a flow rate of 250 nL/min. The data were acquired in data dependent mode with a resolution of 120,000 full-width half maximum at m/z 200 in the survey scan (375-1500 m/z) and with EASY-IC using the reagent ion source (202 m/z) for internal calibration. MS/MS spectra were acquired after precursor isolation in the quadrupole with an isolation window of 1.2 Th, dynamic precursor selection (top speed mode) with a fixed duty cycle time of 3 seconds and dynamic precursor exclusion of 60 seconds. Isolated precursor ions were fragmented by CID with a normalised Collision Energy of 35%. Parallelization was enabled and MS/MS spectra were acquired in the linear ion trap for up to 250 ms with an ion target of 4000 in rapid scan mode.

Raw MS data were analysed using Progenesis QI for Proteomics software v3.0 (Non-linear Dynamics). MS/MS spectra were searched against the UniProt Homo Sapiens Reference proteome (retrieved 15/11/2016) using Mascot v2.5.1 (Matrix Science) allowing for a precursor mass tolerance of 10 ppm and a fragment ion tolerance of 0.5 Da. Carbamidomethylation on Cysteines was defined as fixed modification and variable modifications included Deamidation on Asparagine and Glutamine, and

Oxidation on Methionine. The peptide FDR was set at 1% and all peptides with an ion score higher than 20 were imported into Progenesis QIP.

6.3.6 Normalisation and statistical analysis

Proteins that were defined with at least one unique peptide were included in the protein data set for further analysis. Protein abundance values were normalised by centering on the median abundance of the 90% of proteins with the lowest variance across all runs and scaled by MAD (Keilhauer et al. 2015). Missing values were imputed using k NN with $k = 3$. Statistical analysis was performed as described in **Section 2.4**. Since the normalised log abundance values followed a normal distribution, comparisons of abundance were performed using a Welch's t -test with FDR correction using the Benjamini-Hochberg step-up procedure. GO enrichment analysis was performed in R using TopGO using the weight algorithm and a target-background approach, using all proteins with $p < 0.05$ and $\log_2$ fold change > 0.5 or < -0.5 as the foreground list and all proteins identified as the background list. PCA was performed using normalised, imputed abundance and hierarchical clustering using euclidian distance.

Comparisons of modal EV size and total number were performed using a Mann-Whitney U test or Kruskal-Wallis H test, since the distribution of values was non-normal.

GO enrichment analysis was performed using TopGO with a target-background approach using the weight algorithm. Proteins with unadjusted p -value < 0.05 and $\log_2$ fold change > 0.5 in ALS or healthy control samples were used as the target list and all proteins identified in proteomic analysis as the background list.

6.4 Results

6.4.1 EV size distribution and number

The modal size of and total number of EVs isolated per mL starting CSF were compared between ALS patients, healthy controls and patients with mimic conditions. No significant differences were found (total number $p = 0.685$, modal size $p = 0.390$; **Figure 6.1**). Since a deficit in EV release from fibroblasts and iPSC motor neurons from ALS patients carrying the *C9orf72* repeat expansion, modal size and total EV number was also compared for ALS patients with and without a repeat expansion; there was no significant difference (total number $p = 0.284$, modal size $p = 0.619$; **Figure 6.2**).

The overall variation in EVs number per mL CSF was high (CV 30.02%), which was comparable to the longitudinal variation in individuals (6 subjects; median CV 31.46%, IQR 15.72-42.33%). There was no significant correlation of EV number with PR (Pearson's $r = 0.03$, $p = 0.913$) or age at sampling (Pearson's $r = 0.18$, $p = 0.437$).

6.4.2 The CSF EV proteome in ALS

Due to an excessive quantity of missing data (>50%), proteomic data from 7 mimic samples, 4 healthy controls and 8 ALS patients were excluded from subsequent analysis. Since the number of remaining mimic samples (2) was below that suitable for reliable statistical inference, univariate analysis was performed comparing ALS and healthy control subjects (**Table 6.2**).

A total of 1228 proteins were identified and quantified across all samples, with 3.08% missing values. The median combined CV across all subject groups for all proteins

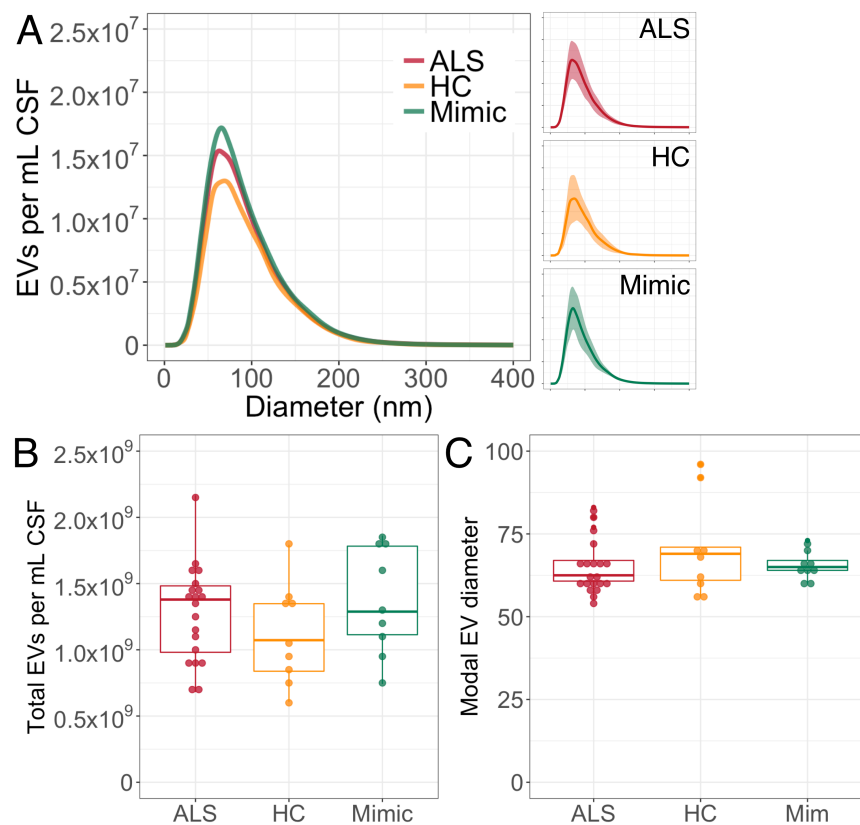


Figure 6.1: **A** Mean size distribution of CSF EVs in patients with ALS, healthy controls and ALS-mimic conditions overlaid. Mean $\pm$ SD for groups shown in neighbouring plots **B** total number of EVs per mL CSF and **C** modal size of extracted EVs. ALS - amyotrophic lateral sclerosis; EV - extracellular vesicle; CSF - cerebrospinal fluid; HC - healthy control; C9 - *C9orf72* hexanucleotide repeat expansion.

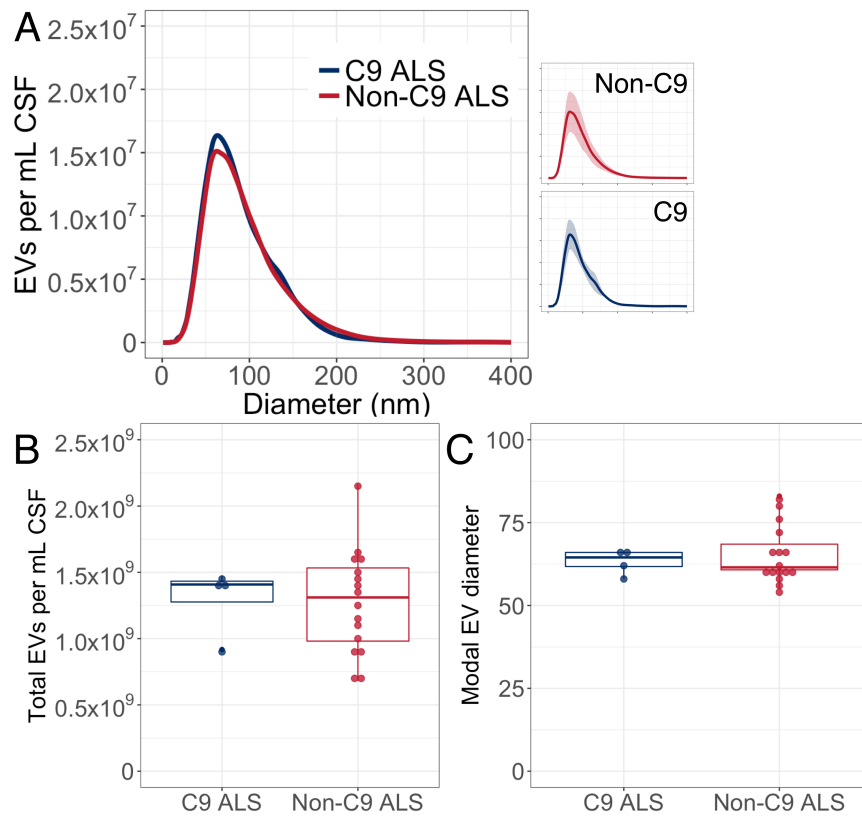


Figure 6.2: **A** Mean size distribution of CSF EVs in patients with *C9*-associated and non-*C9*-associated ALS overlaid. Mean $\pm$ SD for groups shown in neighbouring plots **B** total number of EVs per mL CSF and **C** modal size of extracted EVs. ALS - amyotrophic lateral sclerosis; EV - extracellular vesicle; CSF - cerebrospinal fluid; HC - healthy control; C9 - *C9orf72* hexanucleotide repeat expansion.

	ALS	Mimic conditions	Healthy control	<i>p</i>
n, first visit	12	2	5	-
n, second visit	5	-	-	
n, third visit	3	-	-	
% male	66.7	100	20	0.14*
Age at sampling, mean $\pm$ SD (years)	68.1 $\pm$ 12.0	68.7 $\pm$ 8.6	52.0 $\pm$ 14.8	0.21 ⁺
Age at onset, mean $\pm$ SD (years)	59.7 $\pm$ 12.4	-	-	-
% bulbar onset	50.0	-	-	-
Disease progression rate, median [IQR]	0.53 [0.21-0.63]	-	-	-

Table 6.2: Demographic data for CSF EV proteomic analysis. * Fisher exact test. ⁺Kruskal-Wallis *H* test. ALS – amyotrophic lateral sclerosis; SD - standard deviation; IQR - interquartile range.

was 29.48% (IQR 21.78-40.24%). The median CV for pooled samples was 10.11% (IQR 5.85-16.83%; difference between pooled sample and individual sample CVs $p < 0.001$); pairwise correlation between pool samples was median 0.96 (Pearson's r ; all $p < 0.001$; IQR 0.95-0.97), non-pooled samples pairwise correlation was median 0.86 (IQR 0.78-0.89; difference between pooled sample and individual sample pairwise correlation coefficients $p < 0.001$).

Cross sectional analysis

Multivariate analysis Hierarchical clustering and PCA of first-visit samples did not lead to separation into subject groups (**Figure 6.3**), suggesting that most variance in the data is not attributable to subject group.

Univariate analysis - ALS vs healthy control Comparison of CSF EV protein abundance between ALS patients and healthy controls identified differential abundance of one protein with FDR-adjusted $p < 0.1$, Bleomycin hydrolase (BLMH),

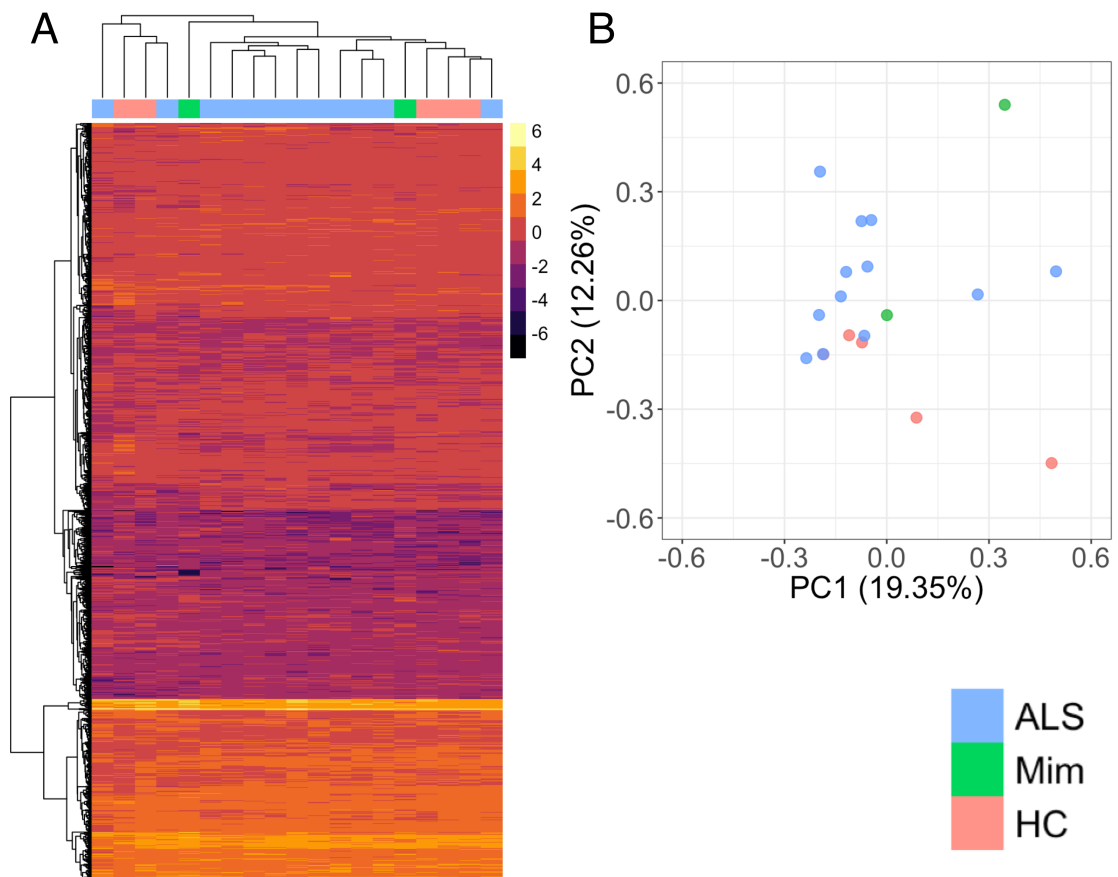


Figure 6.3: **A** Heatmap and hierarchical clustering of protein abundance in the CSF EV proteome. Column colour bar indicates sample group. **B** PCA biplot of CSF EV samples. ALS - amyotrophic lateral sclerosis; EV - extracellular vesicle; CSF - cerebrospinal fluid; HC - healthy control; Mim - mimic conditions; PC - principal component.

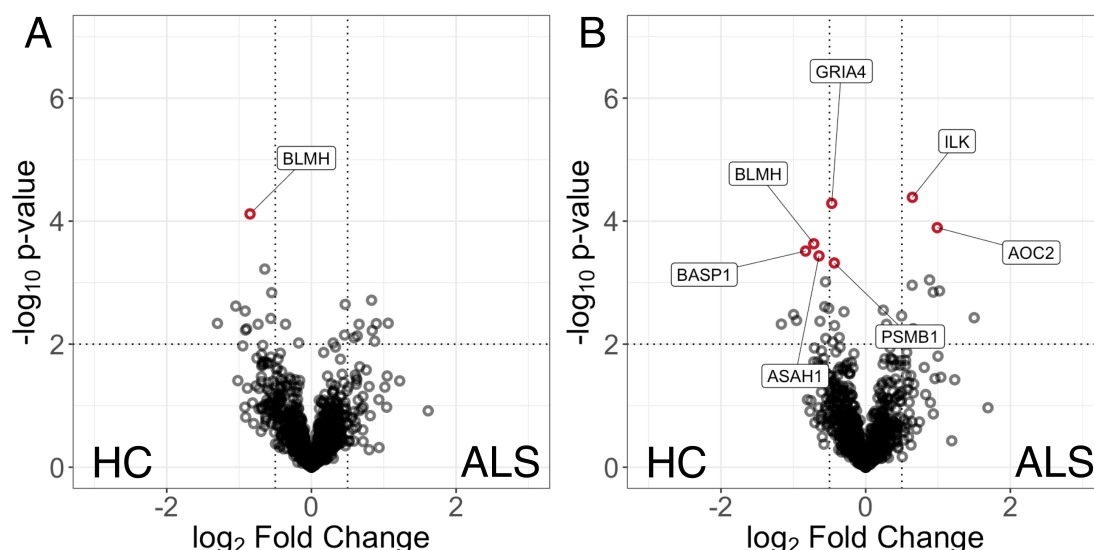


Figure 6.4: Volcano plot of the CSF EV proteome in ALS compared with healthy control subjects. **A** first-visit ALS samples only, **B** samples from all visits for ALS patients. Red points indicate FDR-adjusted $p < 0.1$. Negative log fold change indicates proteins downregulated in ALS compared with controls. Dotted lines indicate $p < 0.001$ and $\log_2$ fold change 0 ± 0.5 . ALS - amyotrophic lateral sclerosis; EV - extracellular vesicle; CSF - cerebrospinal fluid; HC - healthy control; BLMH - Bleomycin hydrolase; GRIA4 - Glutamate receptor 4; BASP1 - Brain acid soluble protein 1; ASAH1 - Acid ceramidase; PSMB1 - Proteasome beta subunit type-1; AOC2 - Retina-specific copper amine oxidase; ILK - Integrin linked protein kinase.

which was downregulated in the CSF EV samples from ALS patients. Including all ALS samples (i.e. those obtained at second and third visits for some subjects), 7 proteins were differentially abundant with FDR-adjusted $p < 0.1$, including down-regulation of BLMH (**Figure 6.4** and **Table 6.3**).

GO enrichment analysis - ALS vs healthy control No GO terms were significant with FDR-adjusted $p < 0.05$. Terms with unadjusted $p < 0.05$ are shown in **Figure 6.5A-B**; the top 10 GO terms ranked by ascending p -value are labelled. Interestingly, the most significant term in proteins upregulated in ALS was "autophagosome" (2 foreground proteins/7 annotated proteins, OR 18.18, $p = 0.005$), whilst several terms in the proteins downregulated in ALS related to the UPS: "proteasome core complex" (3/8 proteins, OR 13.04, $p = 0.001$), "protein polyubiquitination" (4/14 proteins, OR 10.00, $p < 0.001$) and "negative regulation of ubiquitin-protein ligase activity" (3/9 proteins, OR 11.54, $p = 0.002$). Other terms

Protein	ALS vs HC - first visit			ALS vs HC - all visits		
	FC	<i>p</i>	FDR	FC	<i>p</i>	FDR
Integrin linked protein kinase	0.46	0.002	0.364	0.64	<0.001	0.026
Glutamate receptor 4	-0.36	0.005	0.364	-0.47	<0.001	0.026
Retina-specific copper amine oxidase	0.83	0.002	0.364	0.99	<0.001	0.042
Bleomycin hydrolase	-0.84	<0.001	0.087	-0.72	<0.001	0.058
Brain acid soluble protein 1	-0.49	0.026	0.628	-0.83	<0.001	0.058
Acid ceramidase	-0.56	0.004	0.364	-0.65	<0.001	0.058
Proteasome subunit beta type-1	-0.43	0.014	0.518	-0.44	<0.001	0.062

Table 6.3: Relative abundance CSF EV proteins differentially abundant comparing ALS and healthy control samples, using samples from the first-visit only or all samples. ALS – amyotrophic lateral sclerosis; PLS – primary lateral sclerosis; FC – $\log_2$ fold change; *p* – unadjusted *p*-value; FDR – false discovery rate.

of relevance to ALS were identified ("intermediate filament cytoskeleton" in ALS downregulated proteins [6/42 proteins, OR 5.08, $p < 0.001$], "cytosolic small ribosomal subunit" in ALS upregulated proteins [1/1 proteins, OR 50.00, $p = 0.016$]), as well as terms not clearly related to ALS such as "keratinization" (proteins downregulated in ALS, 4/10 proteins, OR 14.29, $p < 0.001$, 4/10 proteins).

Further exploring GO terms, although none reached the FDR threshold in first-visit ALS samples, all proteins in the proteomic dataset encoded by genes annotated to the cellular component term "proteasome core complex" (proteasomal alpha subunits PSMA5, PSMA6, PSMA7 and beta subunits PSMB1, PSMB2, PSMB3, PSMB4, PSMB6), have negative log fold change in ALS compared with healthy controls (**Figure 6.5C**). Of proteins encoded by genes annotated to the biological process term "protein polyubiquitination", (proteasome core complex proteins plus Ubiquitin B (UBB), BLMH, Checkpoint with forkhead and ring finger domains (CHFR), HECT, UBA and WWE domain containing 1 E3 ubiquitin protein ligase (HUWE1), Puromycin-sensitive aminopeptidase (NPEPPS) and E3 ubiquitin-protein ligase RNF213 (RNF213)), only CHFR and RNF213 had positive log fold change in ALS.

Proteins encoded by genes annotated to "autophagosome", though again not meet-

ing the FDR threshold in first-visit samples have predominantly positive log fold change in ALS (FYVE and coiled-coil domain-containing protein 1 (FYCO1), Vacuole membrane protein 1 (VMP1), Sushi repeat-containing protein SRPX2 (SRPX), Ras-related protein Rab-7a (RAB7A) and LAMP2), though three had negative log fold change (Ras-related protein Ral-B (RALB), Ferritin light chain (FTL) and SRPX; **Figure 6.5D**).

Univariate analysis - *C9* ALS vs non-*C9* ALS Given findings of altered EV secretion from iPSC-derived motor neurons (Aoki et al. 2017), differences in the CSF EV proteome in ALS associated with the *C9orf72* hexanucleotide repeat were sought by a comparison of protein abundances between those carrying a *C9* mutation ($n = 3$) and those without ($n = 9$). Six proteins were differentially abundant with FDR-adjusted $p < 0.1$: Transmembrane glycoprotein NMB (GP-NMB), Annexin 11 (ANXA11), Ubiquitin-like modifier-activating enzyme 1 (UBA1), Cytochrome b-245 heavy chain (CYBB), Cofilin-1 (CFL1) and Protein-glutamine gamma-glutamyltransferase 2 (TGM2) (**Figure 6.6A**). Only UBA1, a protein involved with ubiquitination of proteins for degradation by the UPS, had a $\log_2$ fold change of $> \pm 0.5$.

GO enrichment analysis Again, no GO terms were enriched with FDR-adjusted $p < 0.05$. GO terms with unadjusted $p < 0.05$ are shown in **Figure 6.6B-C**. Most of GO terms considered significant at this threshold relate to a single significant protein, including those of most relevance to the pathological processes thought to underlie ALS, such as "cytoplasmic mRNA processing body assembly" (1/1 proteins, OR 100.00, $p = 0.011$) and "ubiquitin activating enzyme activity" (1/1 proteins, OR 100.00, $p = 0.008$). The exceptions to this were "positive regulation of lymphocyte proliferation", upregulated in *C9orf72* ALS, "intermediate filament" (4/42 proteins, OR 6.70, $p < 0.001$), "transferase activity, transferring pentosyl groups" (2/3 proteins, OR 50.00, $p = 0.012$) and "structural molecule activity" (4/110 proteins, OR

2.90, $p = 0.041$).

6.5 Discussion

The experiments described in this chapter explored CSF EVs as they relate to ALS in terms of size distribution and protein cargo. The study of size distribution using NTA did not demonstrate differences in the total number or modal size of EVs purified from the CSF of patients with ALS (including when subdivided by *C9* status). This is in contrast to the *in vitro* findings of Aoki et al. (2017), whereby a significant deficit in EV production from fibroblasts and iPSC motor neurons derived from patients with a *C9orf72* hexanucleotide repeat expansion. This may be a reflection of the multiple cell types that contribute to the CSF EV population (such as oligodendrocytes and choroid plexus as established in **Chapter 5**), or might suggest that the mechanisms leading to altered EV secretion in *C9orf72* expansion-carrying cells *in vitro* do not accurately recapitulate events *in vivo*.

The high variability of the CSF EV concentration was also notable, varying from 0.5 to 2.2×10^9 per mL CSF. Although this could lead to reduced power to detect a significant difference, the median total EV number per mL CSF was extremely similar between ALS and control groups. There was also no correlation with PR or with age (a decline in CSF EV concentration with increasing age was noted in a previous study; Tietje et al. 2014).

Analysis of the EV proteome yielded good proteomic depth, with over 1200 protein groups identified, though the list of candidate markers identified that were differentially abundant between ALS and healthy controls, using a FDR threshold of 10% and $\log_2$ fold change of ± 0.5 was small: just one protein, BLMH, when considering only first-visit ALS samples. BLMH forms a hexameric barrel-shaped protease somewhat homologous to the proteasome (O’Farrell et al. 1999), with diverse proteolytic

activity (Zheng et al. 1998). BLMH has been implicated in Huntington’s disease, through cleavage of huntingtin (Ratovitski et al. 2011), and in Alzheimer’s disease through the processing of amyloid precursor protein (Lefterov et al. 2000). Although clearly requiring validation, BLMH has biological plausibility as a biomarker candidate.

Incorporating data from subsequent visits increased the number of significant proteins to seven, though this likely violates the assumption of independence underlying the statistical analysis techniques and may falsely depress the associated p -values.

There are several potential reasons for this low number of identified candidates. The relatively high variability observed (median combined CV across the cohort of $>29\%$), which is accentuated in low-abundance proteins (median CV of the lowest intensity 10% of proteins 50.06%, IQR 39.64-63.18%) will have contributed, along with the small sample size due to exclusion of samples with large quantities of missing data.

The main signal emerging from GO enrichment analysis, though not meeting a 5% FDR threshold, suggested that involvement of proteostatic mechanisms was detectable in CSF EVs in ALS, with proteins involved in autophagy upregulated in ALS and proteasomal proteins downregulated. Protein homeostasis is implicated in the pathogenesis of ALS (and other neurodegenerative diseases; see **Section 1.2.6**), with mutations in genes encoding both UPS and autophagy proteins associated with the development of ALS.

There is pathological evidence of reduced proteasomal activity and decreases in 20S $\beta 5$ proteasomal subunit levels in spinal cord, along with reduced staining for 20S α subunits in ventral horn motor neurons of patients with ALS (though reductions in the levels of total 20S $\beta 3$ or 19S subunits measured by immunoblot were not found in this study; Kabashi et al. 2012); this is somewhat controversial, as other work has demonstrated increased proteasomal staining in ALS (Mendonca et al. 2006).

Upregulation of autophagy is implied by p62 immunopositivity of neuronal and glial inclusions (Arai et al. 2003; Mizuno et al. 2006). Autophagosomes, though absent from spinal motor neurons in non-ALS pathological specimens, have been observed in both normal-appearing and degenerating spinal motor neurons in ALS, implying upregulation of autophagy (Sasaki 2011).

Although analysis comparing *C9orf72*-expansion associated and non-*C9* ALS yielded more significant proteins at the 5% FDR level, these were of low fold change and GO enrichment analysis demonstrated limited systematic changes in the proteome.

Though the findings of this chapter are partially corroborated by previously published pathological data, further validation is required, particularly given the degree of uncertainty associated with this analysis. Detection of similar changes in CSF EVs from an independent cohort would help to prove that these findings are not spurious. Given the multiplex nature and high sensitivity required, targeted proteomic analysis such as PRM would be well-suited to such a task. In addition, alterations in proteasomal and autophagy-association protein levels in whole CSF should be sought, since at present EV extraction methods remain too complicated and require too much CSF for clinical translation.

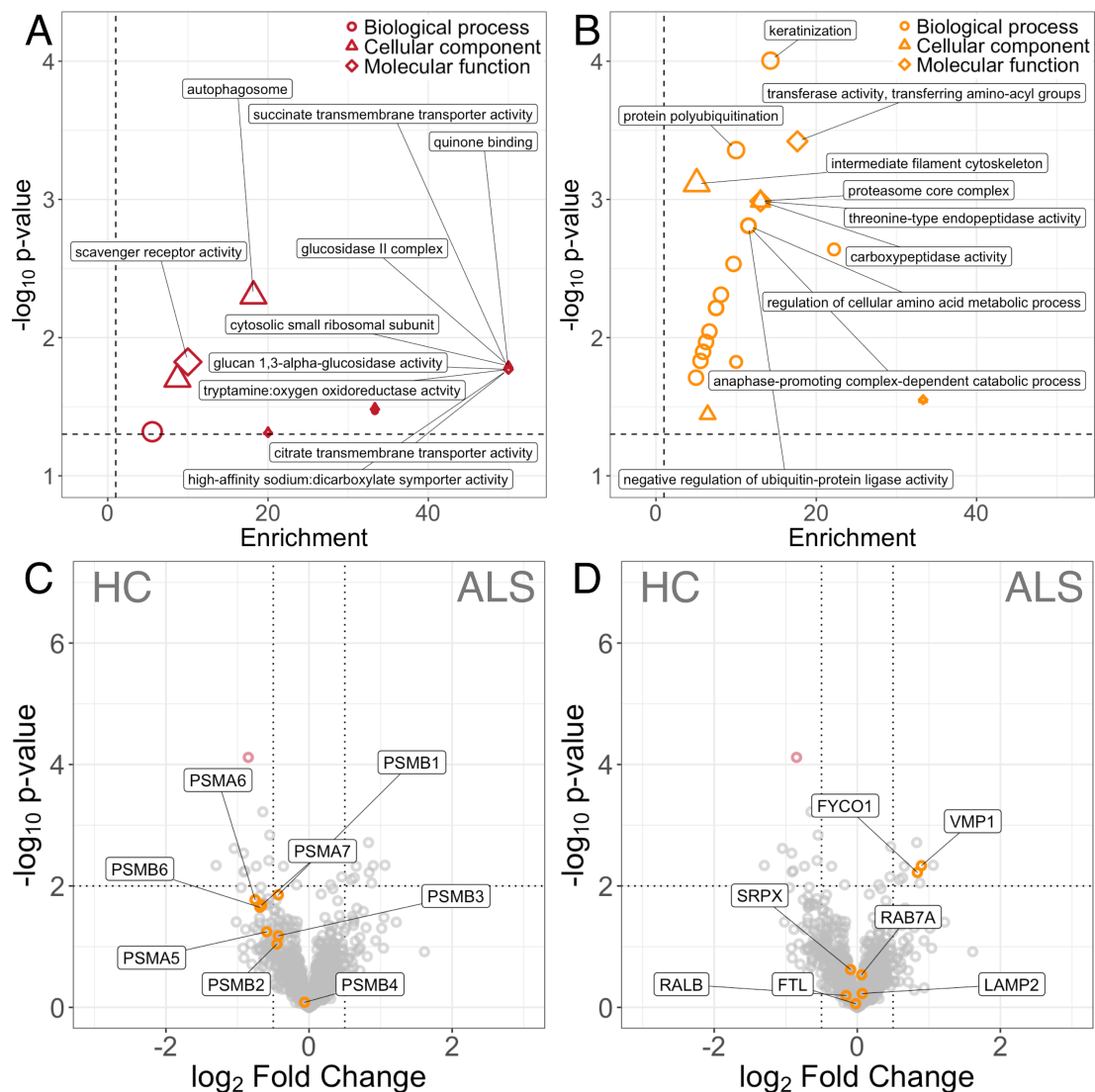


Figure 6.5: GO enrichment analysis of the CSF EV proteome in ALS. **A** GO enrichment analysis of proteins upregulated in ALS $p < 0.05$ and log fold change > 0.5 compared with healthy controls and **B** proteins downregulated in ALS, $p < 0.05$ and fold change < -0.5 compared with healthy controls. Dotted lines indicate $p = 0.05$ and OR 1. Point size proportional to number of foreground proteins annotated to each GO term. Top 10 GO terms in up- and downregulated proteins are labelled. **C** proteins annotated to "proteasome core complex" and **D** "autophagosome" labelled and highlighted with yellow points on volcano plots of first-visit ALS samples vs healthy controls (as in **Figure 6.4A**). ALS - amyotrophic lateral sclerosis; EV - extracellular vesicle; CSF - cerebrospinal fluid; GO - gene ontology; OR - odds ratio.

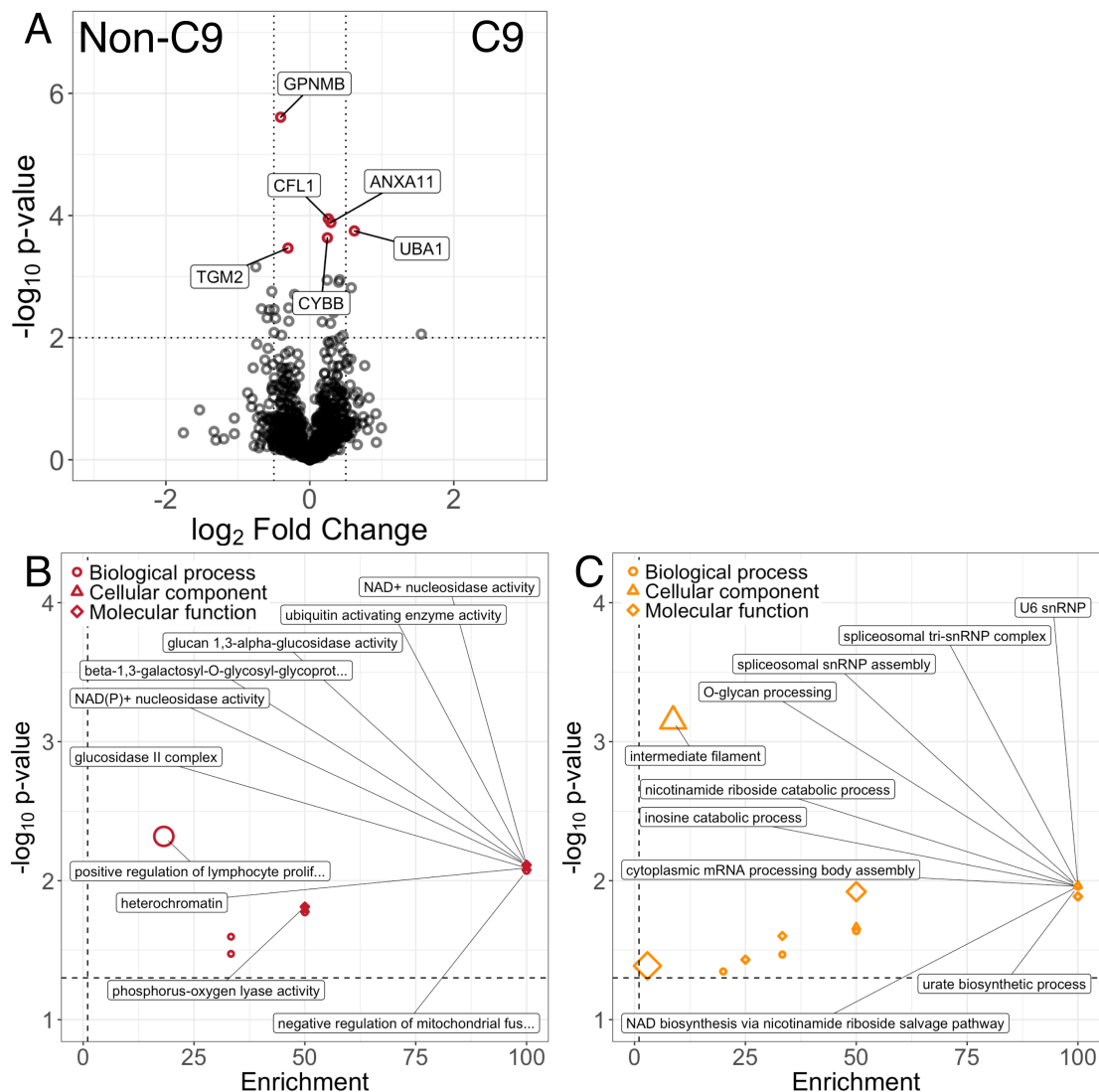


Figure 6.6: **A** Volcano plot of the CSF EV proteome in *C9* ALS compared with non-*C9* ALS subjects (first-visit ALS samples only). Red points indicate FDR-adjusted $p < 0.1$. Negative log fold change indicates proteins downregulated in *C9* ALS compared with controls non-*C9* ALS. Dotted lines indicate $p < 0.001$ and $\log_2$ fold change 0 ± 0.5 . **B-C** GO enrichment analysis of the CSF EV proteome in *C9* ALS. **B** GO enrichment analysis of proteins upregulated in *C9* ALS $p < 0.05$ and $\log_2$ fold change > 0.5 compared non-*C9* ALS and **C** proteins downregulated in *C9* ALS, $p < 0.05$ and fold change < -0.5 compared with non-*C9* ALS. Dotted lines indicate $p = 0.05$ and OR 1. Point size proportional to number of foreground proteins annotated to each GO term. Top 10 GO terms in up- and downregulated proteins are labelled. ALS - amyotrophic lateral sclerosis; EV - extracellular vesicle; CSF - cerebrospinal fluid; GO - gene ontology; OR - odds ratio.

7. Discussion

The aim of the studies described in this thesis was to advance the development of biomarkers of ALS. My novel contributions to this field have been:

1. Development of a deeply-phenotyped cohort of CSF from a range of ALS patients longitudinally, plus comparator disease and healthy control cohorts
2. Proteomic analysis of whole CSF and ELISA-based analysis of emerging candidates
3. Development of an optimised method for CSF EV isolation
4. Comparative proteomic analysis of CSF EVs

This chapter will discuss my results in the wider context of ALS research and outline suggestions for the direction of future research.

7.1 The ALS CSF proteome and the chitinase proteins

The whole CSF proteomic analysis conducted in **Chapter 3** identified elevation in a group of three related glial proteins, the chitinase proteins, in the CSF of ALS patients compared with disease and healthy controls. Chitinase proteins, particularly CHIT1, were found to correlate with the intrinsic aggressiveness of the disease as measured by the PR and survival, though correlations were less close than those observed for neurofilament proteins. These results are very much in keeping with the literature, confirming the findings of one previous proteomic study, similarly identifying elevation of the three proteins, and two ELISA-based studies showing

elevation of CHIT1 (Varghese et al. 2013; Steinacker et al. 2018; Chen et al. 2016b).

Chapters 3 and **4** have also contributed to our understanding of the relationship between the chitinase proteins and clinical dimensions of ALS. Although the association of PR and survival with CHIT1 has previously been demonstrated (Steinacker et al. 2018), this work highlights more complex associations of the three chitinase proteins - CHI3L1 being more closely associated with indices of central dysfunction (i.e. the UMN score and degree of cognitive impairment), and CHIT1 and CHI3L2 being more closely associated with PR.

These differences might be hypothesised to relate to differential expression of chitinase proteins by different glial populations, as is suggested by a combination of existing *post mortem* studies in both ALS and other diseases (Hinsinger et al. 2015; Querol-Vilaseca et al. 2017; Steinacker et al. 2018), or perhaps to different activation states of microglia in different CNS regions. A true understanding of the source of the different CSF chitinases, however, would require a more detailed exploration of chitinase expression by different cell types in *post mortem* tissue from patients with ALS from different CNS regions, preferably correlated with *post mortem* and *ex vivo* CSF levels.

The assessment of diagnostic performance of the chitinase proteins compared with pNFH performed in **Chapters 3** and **4**, combined with the literature comparing CHIT1 with neurofilament proteins in diagnostic performance, suggest that the chitinase proteins are unlikely to overtake neurofilaments as the front-running diagnostic and prognostic biomarkers of ALS. It is possible that, after further study, CSF chitinase levels might form part of a diagnostic panel or be used to refine prognostic estimates for the purposes of clinical trials, but the data presented in **Chapters 3** and **4** do not provide strong support for this.

Although the chitinase proteins do not fulfill the criteria of the "ideal" biomarker of ALS laid out in **Section 1.3** (Turner et al. 2009), they may still prove useful in

future. A worthwhile role for chitinase proteins is likely to involve measuring target engagement and stratification in clinical trials of drugs targeting glial mechanisms, in particular microglial activity. Should such trials prove fruitful, there is a plausible clinical use for chitinase proteins as a theranostic marker, predicting response and permitting tailored ALS therapy. The lack of truly prospective diagnostic cohort in this study (in common with previous studies), the overlap between the cohorts used in the ELISA and proteomic studies, and the modest size of the cohort, particularly the mimic group, are significant limitations of this study which impede extrapolation of the results and necessitate further validation.

Beyond identifying a set of candidate markers, proteomic analysis of the whole CSF proteome suggested that the most florid evidence of ALS detectable in CSF is that of microglial (and possibly astroglial) activity, rather than degeneration of neurons. This could be misconstrued as evidence of the importance of non-autonomous mechanisms in ALS (outlined in **Section 1.2.6**) and the relevance (or even primacy) of the microglial response in its pathogenesis. Activated microglia are secretory cells, releasing multiple cytokines as well as chitinase proteins (Liu et al. 2018; Kanneganti et al. 2012; Li et al. 2018; Kunz et al. 2015), whereas CSF neurofilaments are not secreted. As elucidated in **Chapter 6**, proteomic analysis of whole CSF mostly identifies secreted proteins, meaning that alterations in the proteome are more likely to reflect processes involving secreted factors.

Commensurate with this, the regression analysis performed in **Chapters 3** and **4** suggesting that the association of survival and PR with pNFH is closer than the association of survival and PR with CHIT1, provides supportive evidence for the relative importance of neurodegeneration over microglial activity in relation to the severity of ALS. This does not inform the debate as regards the relevance of microglia to the pathogenesis of ALS; nor does it exclude the possibility that pharmacological modification of the microglial response might be beneficial.

7.2 The CSF EV proteome

In keeping with the general hypothesis of **Chapter 5**, the CSF EV proteome demonstrated improved access to the intracellular proteome when compared to the study of whole CSF and, confirming previous proteomic analysis of CSF EVs, more unconventionally secreted proteins in EVs (Chiasserini et al. 2014). Intriguingly, analysis using topographical mRNA expression data identified differences in the contribution of different brain regions to the whole CSF and CSF EV proteomes, with the cortex contributing more to the whole CSF proteome and the choroid plexus and white matter contributing more to the EV proteome. These findings were partially recapitulated using individual proteomic datasets, which showed overrepresentation of myelin proteins, proteins identified in frontal cortical neurons and hippocampal neurons in the CSF EV-enriched proteome.

Together, these findings are encouraging for CSF EVs as a biomarker source in neurodegenerative diseases, including ALS. The results described in **Chapter 5** suggest that the EV proteome better reflects the intracellular proteome and is therefore more likely to provide markers of events leading to or indicative of neuronal death and linked to the pathogenesis of ALS (such as defective RNA biology or proteostatic mechanisms, described in **Section 1.2.6**).

It is somewhat discouraging, however, that only a small number of current candidate markers of neurodegenerative diseases were identified in the EV-enriched proteome, and that the greatest contribution to the CSF EV proteome appears to arise from white matter tracts and the choroid plexus. Whilst these regions might provide useful biomarkers of ALS (indeed, neurofilament proteins probably represent axonal loss and TDP-43 inclusions are found in oligodendrocytes; Menke et al. 2015; Brettschneider et al. 2014), the most likely source of pathology-specific biomarkers of ALS is grey matter, specifically pyramidal neurons of the agranular motor cortex and motor neurons of the spinal cord, since this is focus of the primary pathologi-

cal and clinical changes in ALS. If choroid plexus and white matter are the major contributors to the CSF EV population, it would have an impact on signal-to-noise ratio and the likelihood of detecting alterations in protein abundance.

Indeed, the analysis of the CSF EV proteome in ALS detected only small changes in protein abundance between ALS and healthy control samples and demonstrated only one *bona fide* significant hit at the 10% FDR threshold, albeit one with a link to proteostasis. This lack of significant hits is likely to be a reflection of the heterogeneity of the disease combined with variability introduced by such a complex extraction protocol and the very small amount of protein available for proteomic analysis.

An agnostic examination of the proteomic data, using GO enrichment highlighted alterations proteostatic mechanisms (most convincingly the proteasome), pathways that are implicated in ALS pathogenesis and with evidence of histopathological alterations in ALS (**Section 1.2.6**). The small numbers studied and lack of validation are major weaknesses that prevent the drawing of firm conclusions from these data. Although proteasomal dysfunction is implicated across neurodegenerative diseases, altered levels of CSF EV proteasomal subunits or BLMH might be helpful in distinguishing ALS from mimic conditions and would be an attractive target to study in people carrying ALS-associated mutations as a potential pre-degenerative marker.

7.3 Future directions

There are a wide range of perceived potential roles for biomarkers in ALS (**Section 1.3**); as yet no biomarker satisfactorily fulfills these roles to the point of widespread clinical use. The neurofilament proteins are currently the front-runners in terms of diagnostic accuracy and prediction of prognosis. Although the neurofilaments have already been studied extensively, estimates of diagnostic accuracy might be

further refined by wider agreement about what constitute ALS mimicking conditions, although this is likely to vary with local practice. An alternative strategy might see clinical adoption of CSF neurofilament levels in the absence of definitive agreement about their diagnostic performance, with refinement after a period of use.

The accuracy of neurofilaments in diagnosing ALS might also involve the use of multiple biomarkers to form a panel, for example combining biomarkers representing different facets of the underlying pathology. This has been successfully performed in Alzheimer's disease, where a ratio of amyloid β (low levels indicating accumulation of amyloid plaques) and tau (a measure of neurodegeneration) outperforms either marker individually (Lewczuk et al. 2018). The assessment described in **Chapters 3 and 4** using a combination of chitinase proteins with pNFH provided somewhat contradictory results, but warrants further appraisal.

Alternatively, other combinations might prove more fruitful - it is notable that the incorporation of amyloid β represents a relatively pathology-specific marker in the Alzheimer's field; using CSF TDP-43 in combination with pNFH in ALS has not been examined, though given that most CSF TDP-43 is blood-derived, this seems unlikely to yield a highly specific marker. Inclusion of a broader panel of markers (such as the cytokine panel suggested by Mitchell et al. (2009), or incorporating multimodal biomarkers such as imaging or neurophysiological data) is an alternative strategy for improving diagnostic accuracy that has not been fully explored.

It is as yet unclear whether including biomarkers in prognostic estimates improves their accuracy compared with clinical measures alone, probably due to the availability of large-scale clinical datasets that enable construction of survival models and relative paucity of accompanying biomarker data (Westeneng et al. 2018). Assessment of survival models with and without biochemical data is the only way that this can be achieved.

There are two other applications of biomarkers that are of increasing urgency. The

first of these is as surrogate markers of clinical effect in drug trials. An effective biomarker of effect would allow a hastening of go-no-go decisions and reduce the cost of clinical trials and the burden on patients who may undergo unnecessarily prolonged treatment with associated side effects (Turner et al. 2015). Neurofilaments are, again, probably the most likely candidates for this (though urinary p75 neurotrophin receptor extracellular domain (p75NTR^{ECD}) may be an alternative; Shepheard et al. 2014). Although CSF neurofilament levels have been shown to decrease following treatment with effective drugs in multiple sclerosis (Gunnarsson et al. 2011), a similar effect has not been proven in ALS despite there being a treatment with demonstrable (albeit modest) disease-modifying efficacy. CSF extraction and biomarker analysis is increasingly incorporated into clinical trials in ALS, but a study specifically looking at the effect of riluzole on longitudinal CSF biomarker levels looks unavoidable. The limited effectiveness of riluzole and concerns about when its effect is mediated (with some evidence to suggest its survival benefit is mediated late in the disease course) are potential issues that might make this problematic and hamper interpretation of a negative result (Fang et al. 2018).

The second of these applications is the prediction of transition from an asymptomatic state to symptomatic ALS (or FTD; so-called phenoconversion) in carriers of ALS-associated gene mutations. The advent of antisense oligonucleotide (ASO) therapies currently under evaluation in *C9orf72* and *SOD1* ALS has rendered this an urgent problem. A negative outcome of ASO therapy trials could be attributed to a number of causes (such as the difficulties associated with delivering large molecules to cells resident deep within the CNS), but major reasons relate to our limited understanding of disease pathogenesis and might reflect the fact that genetic mechanisms are important for the *initiation* of neurodegeneration, but that it is maintained by separate mechanisms. It might therefore be necessary or (even if ASO trials show benefit) favorable to treat mutation carriers prior to the onset of symptoms, even before passing the threshold beyond which symptomatic disease is inevitable (Benatar et al. 2012).

ALS manifests in mutation carriers decades after the motor system is created and refined; using currently available biochemical (CSF neurofilaments) and neurophysiological methods, changes can be detected up to 12 months before symptom onset (Benatar et al. 2018; Vucic et al. 2008; Aggarwal et al. 2002). These changes reflect a short period presymptomatic neurodegeneration (implying that ALS does not follow a gradual erosion of motor neurons), which is therefore likely to be too late to initiate preventative treatment. It is possible that a period of biochemically detectable *tolerated dysfunction* exists that precedes the onset of symptomatic ALS by a longer period and during which development of neurodegeneration might be preventable. There is uncertainty about the existence or duration of such a period, but identifying and defining the biochemical characteristics of this period would be of great value in terms of treatment planning for people carrying ALS associated mutations and would have ramifications in terms of our understanding of the pathomechanisms of ALS in both familial and sporadic disease.

The establishment of appropriate cohorts for this analysis, with long-term longitudinal sampling and phenotyping, is already underway (Benatar et al. 2018). Understanding the longitudinal biochemical changes so far has been confined to the measurement of neurofilaments, which has to-date provided evidence for a short period of detectable axonal degeneration. Testing the levels of the chitinase proteins in this cohort would further advance our knowledge in terms of the temporal relationship between the microglial response and axonal degeneration. Assuming successful validation of BLMH and other CSF EV proteins (in particular related to proteasomal function and autophagy) would also be of great interest, particularly in the period prior to elevation of neurofilaments.

Applying a proteomic methodology to a set of longitudinal samples spanning normal, tolerated dysfunction and phenoconversion timepoints is a highly attractive, albeit challenging, proposition. Given that the whole CSF analytical approach employed in **Chapter 3** detected mainly secreted proteins and did not detect evidence

of neurodegeneration even comparing cases of established ALS with healthy controls, it seems unlikely that it has sufficient depth to identify subtle changes in the intracellular proteome that might be indicative of tolerated dysfunction.

Extensive depletion and prefractionation of whole CSF, which has been successfully used to identify over 3000 proteins in CSF and over 5000 proteins in plasma (Guldbrandsen et al. 2014; Keshishian et al. 2015), perhaps coupled with global analysis of PTMs, would provide a comprehensive analysis of the CSF proteome in great depth, though the sample volume requirement and preanalytical variability are likely to be problematic. Recently developed DIA methods such as SWATH-MS might in future also be used to provide greater proteomic depth (Gillet et al. 2012), though the nascency of this technology limits its use at present.

Increased depth for relevant proteins might alternatively be achieved using the optimised EV extraction method described herein, though the large sample volumes required and complex preanalytical procedure leading to increased variability are considerations that might prevent this. Further optimisation for low sample volumes (such as smaller column volumes or the use of mixed chromatography) could improve upon the sample volume requirements.

Notably, translation of EV biomarkers has proven difficult to-date in all diseases, with no EV protein markers having been adopted for clinical use. This would require major simplification of extraction and detection methods for the clinical diagnostic laboratory; approaches incorporating simultaneous capture and detection of disease-specific EVs or EV proteins based on aptamer or antibody methods show promise and might be customised for different antigens (Yadav et al. 2016; Dong et al. 2018). These methods still require significant refinement before they are ready for clinical use.

7.4 Closing remarks

Effective treatment of ALS could not come quickly enough. Targeted gene therapies including ASOs hold obvious promise for those with ALS linked to monogenic genetic mutations. While significant questions remain about the delivery and timing of such treatments, though these are likely to be tractable problems that might benefit from the use of appropriate biomarkers.

Development of treatment for the majority of so-called sporadic ALS, on the other hand, is a more daunting task. Despite very characteristic clinical features, ALS appears to result from alterations in numerous upstream pathways which converge upon the clinical syndrome of ALS, though their relative contribution is not yet known. A great deal is being learned about the cellular and system changes of ALS through deep phenotyping *in vivo*, *post mortem* tissue, cellular and animal models. Integration of this knowledge into a coherent paradigm that relates these changes to the common downstream clinical and pathological picture of ALS draws ever closer. This provides the tangible hope of effective treatment for ALS and with it a better future for those living with ALS.

8. Publications arising from this research period

8.1 Original articles

A. G. Thompson, E. Gray, M. L. Thezenas *et al.* (2018). “Cerebrospinal fluid macrophage biomarkers in amyotrophic lateral sclerosis”. In: *Ann Neurol* 83.2, pp. 258–268. doi: 10.1002/ana.25143

A. G. Thompson, E. Gray, I. Mager *et al.* (2018). “UFLC-derived CSF extracellular vesicle origin and proteome”. In: *Proteomics* (under review)

M. Salter, E. Corfield . . . **A. G. Thompson** *et al.* (2018). “Initial Identification of a Blood-Based Chromosome Conformation Signature for Aiding in the Diagnosis of Amyotrophic Lateral Sclerosis”. In: *EBioMedicine* 33, pp. 169–184. doi: 10.1016/j.ebiom.2018.06.015

A. G. Thompson, V. Blackwell, R. Marsden *et al.* (2017). “A risk stratifying tool to facilitate safe late-stage percutaneous endoscopic gastrostomy in ALS”. in: *Amyotroph Lateral Scler Frontotemporal Degener* 18.3-4, pp. 243–248. doi: 10.1080/21678421.2016.1274330

H. J. Westeneng, T. A. Debray . . . **A. G. Thompson** *et al.* (2018). “Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model”. In: *Lancet Neurol* 17.5, pp. 423–433. doi: 10.1016/S1474-4422(18)30089-9

8.2 Reviews

A. G. Thompson, E. Gray, S. M. Heman-Ackah, *et al.* (2016). “Extracellular vesicles in neurodegenerative disease - pathogenesis to biomarkers”. In: *Nat Rev Neurol* 12.6, pp. 346–57. doi: 10.1038/nrneurol.2016.68

K. Talbot, E. Feneberg, J. Scaber, **A. G. Thompson**, M. R. Turner (2018). “Amyotrophic lateral sclerosis: the complex path to precision medicine”. In: *J Neurol*. doi: 10.1007/s00415-018-8983-8

P. Lewczuk, P. Riederer . . . **A. G. Thompson** *et al* (2018). “Cerebrospinal fluid and blood biomarkers for neurodegenerative dementias: An update of the Consensus of the Task Force on Biological Markers in Psychiatry of the World Federation of Societies of Biological Psychiatry”. In: *World J Biol Psychiatry* 19.4, pp. 244–328. doi: 10.1080/15622975.2017.1375556

R. Marsden, P. Allan . . . **A. G. Thompson** *et al* (2016). “Nutritional pathway for people with motor neurone disease”. In: *Br J Community Nurs* 21.7, pp. 360–3. doi: 10.12968/bjcn.2016.4 21.7.360

References

- Abbott, N. J. (2004). “Evidence for bulk flow of brain interstitial fluid: significance for physiology and pathology”. In: *Neurochem Int* 45.4, pp. 545–52. DOI: 10.1016/j.neuint.2003.11.006.
- Abrahams, S., J. Newton, E. Niven, J. Foley, and T. H. Bak (2014). “Screening for cognition and behaviour changes in ALS”. In: *Amyotroph Lateral Scler Frontotemporal Degener* 15.1-2, pp. 9–14. DOI: 10.3109/21678421.2013.805784.
- Aebersold, R. and M. Mann (2016). “Mass-spectrometric exploration of proteome structure and function”. In: *Nature* 537.7620, pp. 347–55. DOI: 10.1038/nature19949.
- Aggarwal, A. and G. Nicholson (2002). “Detection of preclinical motor neurone loss in SOD1 mutation carriers using motor unit number estimation”. In: *J Neurol Neurosurg Psychiatry* 73.2, pp. 199–201.
- Alais, S., S. Simoes, D. Baas, S. Lehmann, G. Raposo, J. L. Darlix, and P. Leblanc (2008). “Mouse neuroblastoma cells release prion infectivity associated with exosomal vesicles”. In: *Biol Cell* 100.10, pp. 603–15. DOI: 10.1042/BC20080025.
- Alexianu, M. E., B. K. Ho, A. H. Mohamed, V. La Bella, R. G. Smith, and S. H. Appel (1994). “The role of calcium-binding proteins in selective motoneuron vulnerability in amyotrophic lateral sclerosis”. In: *Ann Neurol* 36.6, pp. 846–58. DOI: 10.1002/ana.410360608.
- Alexianu, M. E., M. Kozovska, and S. H. Appel (2001). “Immune reactivity in a mouse model of familial ALS correlates with disease progression”. In: *Neurology* 57.7, pp. 1282–9.
- Almeida, S. et al. (2013). “Modeling key pathological features of frontotemporal dementia with C9ORF72 repeat expansion in iPSC-derived human neurons”. In: *Acta Neuropathol* 126.3, pp. 385–99. DOI: 10.1007/s00401-013-1149-y.

- Anderson, L. (2012). “Within sight of a rational pipeline for development of protein diagnostics”. In: *Clin Chem* 58.1, pp. 28–30. DOI: 10.1373/clinchem.2011.173377.
- Anderson, L. and C. L. Hunter (2006). “Quantitative mass spectrometric multiple reaction monitoring assays for major plasma proteins”. In: *Mol Cell Proteomics* 5.4, pp. 573–88. DOI: 10.1074/mcp.M500331-MCP200.
- Andreasson, U. et al. (2015). “A Practical Guide to Immunoassay Method Validation”. In: *Front Neurol* 6, p. 179. DOI: 10.3389/fneur.2015.00179.
- Aoki, Y. et al. (2017). “C9orf72 and RAB7L1 regulate vesicle trafficking in amyotrophic lateral sclerosis and frontotemporal dementia”. In: *Brain* 140.4, pp. 887–897. DOI: 10.1093/brain/awx024.
- Arai, T., T. Nonaka, M. Hasegawa, H. Akiyama, M. Yoshida, Y. Hashizume, K. Tsuchiya, T. Oda, and K. Ikeda (2003). “Neuronal and glial inclusions in frontotemporal dementia with or without motor neuron disease are immunopositive for p62”. In: *Neurosci Lett* 342.1-2, pp. 41–4.
- Arnold, E. S. et al. (2013). “ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43”. In: *Proc Natl Acad Sci U S A* 110.8, E736–45. DOI: 10.1073/pnas.1222809110.
- Aronson, J. K. and R. E. Ferner (2017). “Biomarkers-A General Review”. In: *Curr Protoc Pharmacol* 76, pp. 9 23 1-9 23 17. DOI: 10.1002/cpph.19.
- Bakker, L. A., C. D. Schroder, M. A. van Es, P. Westers, J. M. A. Visser-Meily, and L. H. van den Berg (2017). “Assessment of the factorial validity and reliability of the ALSFRS-R: a revision of its measurement model”. In: *J Neurol* 264.7, pp. 1413–1420. DOI: 10.1007/s00415-017-8538-4.
- Barschke, P., P. Oeckl, P. Steinacker, A. Ludolph, and M. Otto (2017). “Proteomic studies in the discovery of cerebrospinal fluid biomarkers for amyotrophic lateral sclerosis”. In: *Expert Rev Proteomics* 14.9, pp. 769–777. DOI: 10.1080/14789450.2017.1365602.

- Baumer, D., D. Hilton, S. M. Paine, M. R. Turner, J. Lowe, K. Talbot, and O. Ansorge (2010). “Juvenile ALS with basophilic inclusions is a FUS proteinopathy with FUS mutations”. In: *Neurology* 75.7, pp. 611–8. DOI: 10.1212/WNL.0b013e3181ed9cde.
- Beers, D. R., W. Zhao, B. Liao, O. Kano, J. Wang, A. Huang, S. H. Appel, and J. S. Henkel (2011). “Neuroinflammation modulates distinct regional and temporal clinical responses in ALS mice”. In: *Brain Behav Immun* 25.5, pp. 1025–35. DOI: 10.1016/j.bbi.2010.12.008.
- Begcevic, I., D. Brinc, A. P. Drabovich, I. Batruch, and E. P. Diamandis (2016). “Identification of brain-enriched proteins in the cerebrospinal fluid proteome by LC-MS/MS profiling and mining of the Human Protein Atlas”. In: *Clin Proteomics* 13, p. 11. DOI: 10.1186/s12014-016-9111-3.
- Bellingham, S. A., B. M. Coleman, and A. F. Hill (2012). “Small RNA deep sequencing reveals a distinct miRNA signature released in exosomes from prion-infected neuronal cells”. In: *Nucleic Acids Res* 40.21, pp. 10937–49. DOI: 10.1093/nar/gks832.
- Benajiba, L. et al. (2009). “TARDBP mutations in motoneuron disease with frontotemporal lobar degeneration”. In: *Ann Neurol* 65.4, pp. 470–3. DOI: 10.1002/ana.21612.
- Benatar, M. and J. Wu (2012). “Presymptomatic studies in ALS: rationale, challenges, and approach”. In: *Neurology* 79.16, pp. 1732–9. DOI: 10.1212/WNL.0b013e31826e9b1d.
- Benatar, M., J. Wu, P. M. Andersen, V. Lombardi, and A. Malaspina (2018). “Neurofilament light: A candidate biomarker of pre-symptomatic ALS and phenocconversion”. In: *Ann Neurol*. DOI: 10.1002/ana.25276.
- Bendtsen, J. D., L. J. Jensen, N. Blom, G. Von Heijne, and S. Brunak (2004). “Feature-based prediction of non-classical and leaderless protein secretion”. In: *Protein Eng Des Sel* 17.4, pp. 349–56. DOI: 10.1093/protein/gzh037.

- Benjamini Yoav; Hochberg, Yosef (1995). “Controlling the false discovery rate: a practical and powerful approach to multiple testing”. In: *Journal of the Royal Statistical Society, Series B* 57.1, pp. 289–300.
- Bensimon, A., A. J. Heck, and R. Aebersold (2012). “Mass spectrometry-based proteomics and network biology”. In: *Annu Rev Biochem* 81, pp. 379–405. DOI: 10.1146/annurev-biochem-072909-100424.
- Biomarkers Definitions Working Group, (2001). “Biomarkers and surrogate endpoints: preferred definitions and conceptual framework”. In: *Clin Pharmacol Ther* 69.3, pp. 89–95. DOI: 10.1067/mcp.2001.113989.
- Bjorkoy, G., T. Lamark, A. Brech, H. Outzen, M. Perander, A. Overvatn, H. Stenmark, and T. Johansen (2005). “p62/SQSTM1 forms protein aggregates degraded by autophagy and has a protective effect on huntingtin-induced cell death”. In: *J Cell Biol* 171.4, pp. 603–14. DOI: 10.1083/jcb.200507002.
- Blasco, H. et al. (2013). “Metabolomics in cerebrospinal fluid of patients with amyotrophic lateral sclerosis: an untargeted approach via high-resolution mass spectrometry”. In: *J Proteome Res* 12.8, pp. 3746–54. DOI: 10.1021/pr400376e.
- Blitterswijk, M. van et al. (2012). “Evidence for an oligogenic basis of amyotrophic lateral sclerosis”. In: *Hum Mol Genet* 21.17, pp. 3776–84. DOI: 10.1093/hmg/dds199.
- Boillee, S., K. Yamanaka, C. S. Lobsiger, N. G. Copeland, N. A. Jenkins, G. Kassiotis, G. Kollias, and D. W. Cleveland (2006). “Onset and progression in inherited ALS determined by motor neurons and microglia”. In: *Science* 312.5778, pp. 1389–92. DOI: 10.1126/science.1123511.
- Bolstad, B. M., R. A. Irizarry, M. Astrand, and T. P. Speed (2003). “A comparison of normalization methods for high density oligonucleotide array data based on variance and bias”. In: *Bioinformatics* 19.2, pp. 185–93.
- Boot, R. G., G. H. Renkema, M. Verhoek, A. Strijland, J. Blik, T. M. de Meulemeester, M. M. Mannens, and J. M. Aerts (1998). “The human chitotriosidase

- gene. Nature of inherited enzyme deficiency”. In: *J Biol Chem* 273.40, pp. 25680–5.
- Bowser, R., M. R. Turner, and J. Shefner (2011). “Biomarkers in amyotrophic lateral sclerosis: opportunities and limitations”. In: *Nat Rev Neurol* 7.11, pp. 631–8. DOI: 10.1038/nrneuro1.2011.151.
- Boylan, K. B. et al. (2013). “Phosphorylated neurofilament heavy subunit (pNF-H) in peripheral blood and CSF as a potential prognostic biomarker in amyotrophic lateral sclerosis”. In: *J Neurol Neurosurg Psychiatry* 84.4, pp. 467–72. DOI: 10.1136/jnnp-2012-303768.
- Boylan, K., C. Yang, J. Crook, K. Overstreet, M. Heckman, Y. Wang, D. Borchelt, and G. Shaw (2009). “Immunoreactivity of the phosphorylated axonal neurofilament H subunit (pNF-H) in blood of ALS model rodents and ALS patients: evaluation of blood pNF-H as a potential ALS biomarker”. In: *J Neurochem* 111.5, pp. 1182–91. DOI: 10.1111/j.1471-4159.2009.06386.x.
- Braak, H., J. Brettschneider, A. C. Ludolph, V. M. Lee, J. Q. Trojanowski, and K. Del Tredici (2013). “Amyotrophic lateral sclerosis—a model of corticofugal axonal spread”. In: *Nat Rev Neurol* 9.12, pp. 708–14. DOI: 10.1038/nrneuro1.2013.221.
- Brettschneider, J., H. Mogel, V. Lehmensiek, T. Ahlert, S. Sussmuth, A. C. Ludolph, and H. Tumani (2008). “Proteome analysis of cerebrospinal fluid in amyotrophic lateral sclerosis (ALS)”. In: *Neurochem Res* 33.11, pp. 2358–63. DOI: 10.1007/s11064-008-9742-5.
- Brettschneider, J., A. Petzold, S. D. Sussmuth, A. C. Ludolph, and H. Tumani (2006). “Axonal damage markers in cerebrospinal fluid are increased in ALS”. In: *Neurology* 66.6, pp. 852–6. DOI: 10.1212/01.wnl.0000203120.85850.54.
- Brettschneider, J., J. B. Toledo, V. M. Van Deerlin, L. Elman, L. McCluskey, V. M. Lee, and J. Q. Trojanowski (2012). “Microglial activation correlates with disease progression and upper motor neuron clinical symptoms in amyotrophic lateral sclerosis”. In: *PLoS One* 7.6, e39216. DOI: 10.1371/journal.pone.0039216.

- Brettschneider, J. et al. (2013). “Stages of pTDP-43 pathology in amyotrophic lateral sclerosis”. In: *Ann Neurol* 74.1, pp. 20–38. DOI: 10.1002/ana.23937.
- Brettschneider, J. et al. (2014). “TDP-43 pathology and neuronal loss in amyotrophic lateral sclerosis spinal cord”. In: *Acta Neuropathol* 128.3, pp. 423–37. DOI: 10.1007/s00401-014-1299-6.
- Brinker, T., E. Stopa, J. Morrison, and P. Klinge (2014). “A new look at cerebrospinal fluid circulation”. In: *Fluids Barriers CNS* 11, p. 10. DOI: 10.1186/2045-8118-11-10.
- Brooks, B. R. (1994). “El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial "Clinical limits of amyotrophic lateral sclerosis" workshop contributors”. In: *J Neurol Sci* 124 Suppl, pp. 96–107.
- Brooks, B. R., R. G. Miller, M. Swash, T. L. Munsat, and Diseases World Federation of Neurology Research Group on Motor Neuron (2000a). “El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis”. In: *Amyotroph Lateral Scler Other Motor Neuron Disord* 1.5, pp. 293–9.
- (2000b). “El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis”. In: *Amyotroph Lateral Scler Other Motor Neuron Disord* 1.5, pp. 293–9.
- Brown R. H., Jr. and A. Al-Chalabi (2017). “Amyotrophic Lateral Sclerosis”. In: *N Engl J Med* 377.16, p. 1602. DOI: 10.1056/NEJMc1710379.
- Brugman, F., J. H. Wokke, J. M. Vianney de Jong, H. Franssen, C. G. Faber, and L. H. Van den Berg (2005). “Primary lateral sclerosis as a phenotypic manifestation of familial ALS”. In: *Neurology* 64.10, pp. 1778–9. DOI: 10.1212/01.WNL.0000162033.47893.F7.
- Bruins, A. P. (1998). “Mechanistic aspects of electrospray ionization”. In: *Journal of Chromatography A* 794.1-2, pp. 345–357.

- Bubis, J. A., L. I. Levitsky, M. V. Ivanov, I. A. Tarasova, and M. V. Gorshkov (2017). “Comparative evaluation of label-free quantification methods for shotgun proteomics”. In: *Rapid Commun Mass Spectrom* 31.7, pp. 606–612. DOI: 10.1002/rcm.7829.
- Bury, J. J., J. R. Highley, J. Cooper-Knock, E. F. Goodall, A. Higginbottom, C. J. McDermott, P. G. Ince, P. J. Shaw, and J. Kirby (2016). “Oligogenic inheritance of optineurin (OPTN) and C9ORF72 mutations in ALS highlights localisation of OPTN in the TDP-43-negative inclusions of C9ORF72-ALS”. In: *Neuropathology* 36.2, pp. 125–34. DOI: 10.1111/neup.12240.
- Byrne, S., C. Walsh, C. Lynch, P. Bede, M. Elamin, K. Kenna, R. McLaughlin, and O. Hardiman (2011). “Rate of familial amyotrophic lateral sclerosis: a systematic review and meta-analysis”. In: *J Neurol Neurosurg Psychiatry* 82.6, pp. 623–7. DOI: 10.1136/jnnp.2010.224501.
- Cahoy, J. D. et al. (2008). “A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function”. In: *J Neurosci* 28.1, pp. 264–78. DOI: 10.1523/JNEUROSCI.4178-07.2008.
- Callister, S. J., R. C. Barry, J. N. Adkins, E. T. Johnson, W. J. Qian, B. J. Webb-Robertson, R. D. Smith, and M. S. Lipton (2006). “Normalization approaches for removing systematic biases associated with mass spectrometry and label-free proteomics”. In: *J Proteome Res* 5.2, pp. 277–86. DOI: 10.1021/pr0503001.
- Carvalho, M. D. and M. Swash (2009). “Awaji diagnostic algorithm increases sensitivity of El Escorial criteria for ALS diagnosis”. In: *Amyotroph Lateral Scler* 10.1, pp. 53–7. DOI: 10.1080/17482960802521126.
- Carvalho, M. de et al. (2008). “Electrodiagnostic criteria for diagnosis of ALS”. In: *Clin Neurophysiol* 119.3, pp. 497–503. DOI: 10.1016/j.clinph.2007.09.143.
- Cedarbaum, J. M., N. Stambler, E. Malta, C. Fuller, D. Hilt, B. Thurmond, and A. Nakanishi (1999). “The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III)”. In: *J Neurol Sci* 169.1-2, pp. 13–21.

- Cellura, E., R. Spataro, A. C. Taiello, and V. La Bella (2012). “Factors affecting the diagnostic delay in amyotrophic lateral sclerosis”. In: *Clin Neurol Neurosurg* 114.6, pp. 550–4. DOI: 10.1016/j.clineuro.2011.11.026.
- Al-Chalabi, A., P. M. Andersen, P. Nilsson, B. Chioza, J. L. Andersson, C. Russ, C. E. Shaw, J. F. Powell, and P. N. Leigh (1999). “Deletions of the heavy neurofilament subunit tail in amyotrophic lateral sclerosis”. In: *Hum Mol Genet* 8.2, pp. 157–64.
- Al-Chalabi, A., F. Fang, M. F. Hanby, P. N. Leigh, C. E. Shaw, W. Ye, and F. Rijdsdijk (2010). “An estimate of amyotrophic lateral sclerosis heritability using twin data”. In: *J Neurol Neurosurg Psychiatry* 81.12, pp. 1324–6. DOI: 10.1136/jnnp.2010.207464.
- Al-Chalabi, A. et al. (2014). “Analysis of amyotrophic lateral sclerosis as a multistep process: a population-based modelling study”. In: *Lancet Neurol* 13.11, pp. 1108–1113. DOI: 10.1016/S1474-4422(14)70219-4.
- Chang, L. and M. J. Monteiro (2015). “Defective Proteasome Delivery of Polyubiquitinated Proteins by Ubiquilin-2 Proteins Containing ALS Mutations”. In: *PLoS One* 10.6, e0130162. DOI: 10.1371/journal.pone.0130162.
- Chen, X., Y. Chen, Q. Wei, R. Ou, B. Cao, B. Zhao, and H. F. Shang (2016a). “Assessment of a multiple biomarker panel for diagnosis of amyotrophic lateral sclerosis”. In: *BMC Neurol* 16, p. 173. DOI: 10.1186/s12883-016-0689-x.
- Chen, Y., X. H. Liu, J. J. Wu, H. M. Ren, J. Wang, Z. T. Ding, and Y. P. Jiang (2016b). “Proteomic analysis of cerebrospinal fluid in amyotrophic lateral sclerosis”. In: *Exp Ther Med* 11.6, pp. 2095–2106. DOI: 10.3892/etm.2016.3210.
- Chiasserini, D., J. R. van Weering, S. R. Piersma, T. V. Pham, A. Malekzadeh, C. E. Teunissen, H. de Wit, and C. R. Jimenez (2014). “Proteomic analysis of cerebrospinal fluid extracellular vesicles: a comprehensive dataset”. In: *J Proteomics* 106, pp. 191–204. DOI: 10.1016/j.jprot.2014.04.028.
- Chio, A., A. Calvo, C. Moglia, L. Mazzini, G. Mora, and Parals study group (2011). “Phenotypic heterogeneity of amyotrophic lateral sclerosis: a population based

- study". In: *J Neurol Neurosurg Psychiatry* 82.7, pp. 740–6. DOI: 10.1136/jnnp.2010.235952.
- Chio, A., A. Canosa, S. Gallo, C. Moglia, A. Ilardi, S. Cammarosano, D. Papurello, and A. Calvo (2012). "Pain in amyotrophic lateral sclerosis: a population-based controlled study". In: *Eur J Neurol* 19.4, pp. 551–5. DOI: 10.1111/j.1468-1331.2011.03540.x.
- Chio, A., G. Logroscino, B. J. Traynor, J. Collins, J. C. Simeone, L. A. Goldstein, and L. A. White (2013a). "Global epidemiology of amyotrophic lateral sclerosis: a systematic review of the published literature". In: *Neuroepidemiology* 41.2, pp. 118–30. DOI: 10.1159/000351153.
- Chio, A., B. J. Traynor, F. Lombardo, M. Fimognari, A. Calvo, P. Ghiglione, R. Mutani, and G. Restagno (2008). "Prevalence of SOD1 mutations in the Italian ALS population". In: *Neurology* 70.7, pp. 533–7. DOI: 10.1212/01.wnl.0000299187.90432.3f.
- Chio, A. et al. (2013b). "UNC13A influences survival in Italian amyotrophic lateral sclerosis patients: a population-based study". In: *Neurobiol Aging* 34.1, 357 e1-5. DOI: 10.1016/j.neurobiolaging.2012.07.016.
- Cirulli, E. T. et al. (2015). "Exome sequencing in amyotrophic lateral sclerosis identifies risk genes and pathways". In: *Science* 347.6229, pp. 1436–41. DOI: 10.1126/science.aaa3650.
- Cleary, J. D. and L. P. Ranum (2013). "Repeat-associated non-ATG (RAN) translation in neurological disease". In: *Hum Mol Genet* 22.R1, R45–51. DOI: 10.1093/hmg/ddt371.
- Collins, M. A., J. An, B. L. Hood, T. P. Conrads, and R. P. Bowser (2015). "Label-Free LC-MS/MS Proteomic Analysis of Cerebrospinal Fluid Identifies Protein/Pathway Alterations and Candidate Biomarkers for Amyotrophic Lateral Sclerosis". In: *J Proteome Res* 14.11, pp. 4486–501. DOI: 10.1021/acs.jproteome.5b00804.

- Colombo, M., G. Raposo, and C. Thery (2014). “Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles”. In: *Annu Rev Cell Dev Biol* 30, pp. 255–89. DOI: 10.1146/annurev-cellbio-101512-122326.
- Cooper-Knock, J. et al. (2014). “Sequestration of multiple RNA recognition motif-containing proteins by C9orf72 repeat expansions”. In: *Brain* 137.Pt 7, pp. 2040–51. DOI: 10.1093/brain/awu120.
- Corcia, P., P. F. Pradat, F. Salachas, G. Bruneteau, Nl Forestier, D. Seilhean, J. J. Hauw, and V. Meininger (2008). “Causes of death in a post-mortem series of ALS patients”. In: *Amyotroph Lateral Scler* 9.1, pp. 59–62. DOI: 10.1080/17482960701656940.
- Cottrell, J. S. (2011). “Protein identification using MS/MS data”. In: *J Proteomics* 74.10, pp. 1842–51. DOI: 10.1016/j.jprot.2011.05.014.
- Cox, J., M. Y. Hein, C. A. Lubner, I. Paron, N. Nagaraj, and M. Mann (2014). “Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ”. In: *Mol Cell Proteomics* 13.9, pp. 2513–26. DOI: 10.1074/mcp.M113.031591.
- Cronin, S., O. Hardiman, and B. J. Traynor (2007). “Ethnic variation in the incidence of ALS: a systematic review”. In: *Neurology* 68.13, pp. 1002–7. DOI: 10.1212/01.wnl.0000258551.96893.6f.
- Cudkowicz, M. E., D. McKenna-Yasek, P. E. Sapp, W. Chin, B. Geller, D. L. Hayden, D. A. Schoenfeld, B. A. Hosler, H. R. Horvitz, and R. H. Brown (1997). “Epidemiology of mutations in superoxide dismutase in amyotrophic lateral sclerosis”. In: *Ann Neurol* 41.2, pp. 210–21. DOI: 10.1002/ana.410410212.
- Cudkowicz, M. E. et al. (2010). “Toward more efficient clinical trials for amyotrophic lateral sclerosis”. In: *Amyotroph Lateral Scler* 11.3, pp. 259–65. DOI: 10.3109/17482960903358865.
- Darmanis, S., S. A. Sloan, Y. Zhang, M. Enge, C. Caneda, L. M. Shuer, M. G. Hayden Gephart, B. A. Barres, and S. R. Quake (2015). “A survey of human

- brain transcriptome diversity at the single cell level”. In: *Proc Natl Acad Sci U S A* 112.23, pp. 7285–90. DOI: 10.1073/pnas.1507125112.
- Davis, J., M. Maes, A. Andreazza, J. J. McGrath, S. J. Tye, and M. Berk (2015). “Towards a classification of biomarkers of neuropsychiatric disease: from encompass to compass”. In: *Mol Psychiatry* 20.2, pp. 152–3. DOI: 10.1038/mp.2014.139.
- De Schaepdryver, M., A. Jeromin, B. Gille, K. G. Claeys, V. Herbst, B. Brix, P. Van Damme, and K. Poesen (2017). “Comparison of elevated phosphorylated neurofilament heavy chains in serum and cerebrospinal fluid of patients with amyotrophic lateral sclerosis”. In: *J Neurol Neurosurg Psychiatry*. DOI: 10.1136/jnnp-2017-316605.
- DeLong, E. R., D. M. DeLong, and D. L. Clarke-Pearson (1988). “Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach”. In: *Biometrics* 44.3, pp. 837–45.
- Demartini, DR. (2013). “A Short Overview of the Components in Mass Spectrometry Instrumentation for Proteomics Analyses”. In: *Tandem Mass Spectrometry*. Ed. by J. Prasain. Vol. 1. InTechOpen. Chap. 2. DOI: 10.5772/54484.
- Deng, H. X. et al. (2011). “Mutations in UBQLN2 cause dominant X-linked juvenile and adult-onset ALS and ALS/dementia”. In: *Nature* 477.7363, pp. 211–5. DOI: 10.1038/nature10353.
- Deng, H., K. Gao, and J. Jankovic (2014). “The role of FUS gene variants in neurodegenerative diseases”. In: *Nat Rev Neurol* 10.6, pp. 337–48. DOI: 10.1038/nrneurol.2014.78.
- Deshaies, J. E. et al. (2018). “TDP-43 regulates the alternative splicing of hnRNP A1 to yield an aggregation-prone variant in amyotrophic lateral sclerosis”. In: *Brain*. DOI: 10.1093/brain/awy062.
- Di Rosa, M., G. Distefano, K. Zorena, and L. Malaguarnera (2016). “Chitinases and immunity: Ancestral molecules with new functions”. In: *Immunobiology* 221.3, pp. 399–411. DOI: 10.1016/j.imbio.2015.11.014.

- Diekstra, F. P. et al. (2012). “UNC13A is a modifier of survival in amyotrophic lateral sclerosis”. In: *Neurobiol Aging* 33.3, 630 e3-8. DOI: 10.1016/j.neurobiolaging.2011.10.029.
- Diz, A. P., M. Truebano, and D. O. Skibinski (2009). “The consequences of sample pooling in proteomics: an empirical study”. In: *Electrophoresis* 30.17, pp. 2967–75. DOI: 10.1002/elps.200900210.
- Dong, H., H. Chen, J. Jiang, H. Zhang, C. Cai, and Q. Shen (2018). “Highly Sensitive Electrochemical Detection of Tumor Exosomes Based on Aptamer Recognition-Induced Multi-DNA Release and Cyclic Enzymatic Amplification”. In: *Anal Chem* 90.7, pp. 4507–4513. DOI: 10.1021/acs.analchem.7b04863.
- Donnelly, C. J. et al. (2013). “RNA toxicity from the ALS/FTD C9ORF72 expansion is mitigated by antisense intervention”. In: *Neuron* 80.2, pp. 415–28. DOI: 10.1016/j.neuron.2013.10.015.
- Dormann, D. and C. Haass (2013). “Fused in sarcoma (FUS): an oncogene goes awry in neurodegeneration”. In: *Mol Cell Neurosci* 56, pp. 475–86. DOI: 10.1016/j.mcn.2013.03.006.
- Drummond, E. S., S. Nayak, B. Ueberheide, and T. Wisniewski (2015). “Proteomic analysis of neurons microdissected from formalin-fixed, paraffin-embedded Alzheimer’s disease brain tissue”. In: *Sci Rep* 5, p. 15456. DOI: 10.1038/srep15456.
- Dugas, J. C., Y. C. Tai, T. P. Speed, J. Ngai, and B. A. Barres (2006). “Functional genomic analysis of oligodendrocyte differentiation”. In: *J Neurosci* 26.43, pp. 10967–83. DOI: 10.1523/JNEUROSCI.2572-06.2006.
- Elden, A. C. et al. (2010). “Ataxin-2 intermediate-length polyglutamine expansions are associated with increased risk for ALS”. In: *Nature* 466.7310, pp. 1069–75. DOI: 10.1038/nature09320.
- Elias, J. E. and S. P. Gygi (2007). “Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry”. In: *Nat Methods* 4.3, pp. 207–14. DOI: 10.1038/nmeth1019.

- Elo, L. L., S. Filen, R. Lahesmaa, and T. Aittokallio (2008). “Reproducibility-optimized test statistic for ranking genes in microarray studies”. In: *IEEE/ACM Trans Comput Biol Bioinform* 5.3, pp. 423–31. DOI: 10.1109/tcbb.2007.1078.
- Es, M. A. van et al. (2011). “Angiogenin variants in Parkinson disease and amyotrophic lateral sclerosis”. In: *Ann Neurol* 70.6, pp. 964–73. DOI: 10.1002/ana.22611.
- Fang, T., A. Al Khleifat, J. H. Meurgey, A. Jones, P. N. Leigh, G. Bensimon, and A. Al-Chalabi (2018). “Stage at which riluzole treatment prolongs survival in patients with amyotrophic lateral sclerosis: a retrospective analysis of data from a dose-ranging study”. In: *Lancet Neurol* 17.5, pp. 416–422. DOI: 10.1016/S1474-4422(18)30054-1.
- Fang, Y. S. et al. (2014). “Full-length TDP-43 forms toxic amyloid oligomers that are present in frontotemporal lobar dementia-TDP patients”. In: *Nat Commun* 5, p. 4824. DOI: 10.1038/ncomms5824.
- Fecto, F. et al. (2011). “SQSTM1 mutations in familial and sporadic amyotrophic lateral sclerosis”. In: *Arch Neurol* 68.11, pp. 1440–6. DOI: 10.1001/archneurol.2011.250.
- Feiler, M. S. et al. (2015). “TDP-43 is intercellularly transmitted across axon terminals”. In: *J Cell Biol* 211.4, pp. 897–911. DOI: 10.1083/jcb.201504057.
- Feneberg, E., P. Steinacker, S. Lehnert, A. Schneider, P. Walther, D. R. Thal, M. Linsenmeier, A. C. Ludolph, and M. Otto (2014). “Limited role of free TDP-43 as a diagnostic tool in neurodegenerative diseases”. In: *Amyotroph Lateral Scler Frontotemporal Degener* 15.5-6, pp. 351–6. DOI: 10.3109/21678421.2014.905606.
- Feneberg, E. et al. (2018). “Multicenter evaluation of neurofilaments in early symptom onset amyotrophic lateral sclerosis”. In: *Neurology* 90.1, e22–e30. DOI: 10.1212/WNL.0000000000004761.
- Fenn, J. B., M. Mann, C. K. Meng, S. F. Wong, and C. M. Whitehouse (1989). “Electrospray ionization for mass spectrometry of large biomolecules”. In: *Science* 246.4926, pp. 64–71.

- Fevrier, B., D. Vilette, F. Archer, D. Loew, W. Faigle, M. Vidal, H. Laude, and G. Raposo (2004). “Cells release prions in association with exosomes”. In: *Proc Natl Acad Sci U S A* 101.26, pp. 9683–8. DOI: 10.1073/pnas.0308413101.
- Fiandaca, M. S. et al. (2015). “Identification of preclinical Alzheimer’s disease by a profile of pathogenic proteins in neurally derived blood exosomes: A case-control study”. In: *Alzheimers Dement* 11.6, 600–7 e1. DOI: 10.1016/j.jalz.2014.06.008.
- Freischmidt, A. et al. (2015). “Haploinsufficiency of TBK1 causes familial ALS and fronto-temporal dementia”. In: *Nat Neurosci* 18.5, pp. 631–6. DOI: 10.1038/nn.4000.
- Furukawa, T. et al. (2015). “CSF cytokine profile distinguishes multifocal motor neuropathy from progressive muscular atrophy”. In: *Neurol Neuroimmunol Neuroinflamm* 2.5, e138. DOI: 10.1212/NXI.0000000000000138.
- Furukawa, Y., K. Kaneko, S. Watanabe, K. Yamanaka, and N. Nukina (2011). “A seeding reaction recapitulates intracellular formation of Sarkosyl-insoluble trans-activation response element (TAR) DNA-binding protein-43 inclusions”. In: *J Biol Chem* 286.21, pp. 18664–72. DOI: 10.1074/jbc.M111.231209.
- Gallagher, J. P. (1989). “Pathologic laughter and crying in ALS: a search for their origin”. In: *Acta Neurol Scand* 80.2, pp. 114–7.
- Ganesalingam, J., J. An, C. E. Shaw, G. Shaw, D. Lacomis, and R. Bowser (2011). “Combination of neurofilament heavy chain and complement C3 as CSF biomarkers for ALS”. In: *J Neurochem* 117.3, pp. 528–37. DOI: 10.1111/j.1471-4159.2011.07224.x.
- Garbuzova-Davis, S. and P. R. Sanberg (2014). “Blood-CNS Barrier Impairment in ALS patients versus an animal model”. In: *Front Cell Neurosci* 8, p. 21. DOI: 10.3389/fncel.2014.00021.
- GE Healthcare Life Sciences, (2014). *Size Exclusion Chromatography Principles and Methods*. Pamphlet. Uppsala, Sweden.

- Gerber, S. A., J. Rush, O. Stemman, M. W. Kirschner, and S. P. Gygi (2003). “Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS”. In: *Proc Natl Acad Sci U S A* 100.12, pp. 6940–5. DOI: 10.1073/pnas.0832254100.
- Gijssels, I. et al. (2012). “A C9orf72 promoter repeat expansion in a Flanders-Belgian cohort with disorders of the frontotemporal lobar degeneration-amyotrophic lateral sclerosis spectrum: a gene identification study”. In: *Lancet Neurol* 11.1, pp. 54–65. DOI: 10.1016/S1474-4422(11)70261-7.
- Gillet, L. C., P. Navarro, S. Tate, H. Rost, N. Selevsek, L. Reiter, R. Bonner, and R. Aebersold (2012). “Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis”. In: *Mol Cell Proteomics* 11.6, O111 016717. DOI: 10.1074/mcp.0111.016717.
- Gitcho, M. A. et al. (2008). “TDP-43 A315T mutation in familial motor neuron disease”. In: *Ann Neurol* 63.4, pp. 535–8. DOI: 10.1002/ana.21344.
- Goetzl, E. J., E. L. Abner, G. A. Jicha, D. Kapogiannis, and J. B. Schwartz (2018). “Declining levels of functionally specialized synaptic proteins in plasma neuronal exosomes with progression of Alzheimer’s disease”. In: *FASEB J* 32.2, pp. 888–893. DOI: 10.1096/fj.201700731R.
- Goetzl, E. J., A. Boxer, J. B. Schwartz, E. L. Abner, R. C. Petersen, B. L. Miller, and D. Kapogiannis (2015). “Altered lysosomal proteins in neural-derived plasma exosomes in preclinical Alzheimer disease”. In: *Neurology* 85.1, pp. 40–7. DOI: 10.1212/WNL.0000000000001702.
- Gold, L. et al. (2010). “Aptamer-based multiplexed proteomic technology for biomarker discovery”. In: *PLoS One* 5.12, e15004. DOI: 10.1371/journal.pone.0015004.
- Gomes, C., S. Keller, P. Altevogt, and J. Costa (2007). “Evidence for secretion of Cu,Zn superoxide dismutase via exosomes from a cell model of amyotrophic lateral sclerosis”. In: *Neurosci Lett* 428.1, pp. 43–6. DOI: 10.1016/j.neulet.2007.09.024.

- Gordon, P. H., B. Cheng, I. B. Katz, M. Pinto, A. P. Hays, H. Mitsumoto, and L. P. Rowland (2006). “The natural history of primary lateral sclerosis”. In: *Neurology* 66.5, pp. 647–53. DOI: 10.1212/01.wnl.0000200962.94777.71.
- Gordon, P. H., B. Cheng, F. Salachas, P. F. Pradat, G. Bruneteau, P. Corcia, L. Lacomblez, and V. Meininger (2010). “Progression in ALS is not linear but is curvilinear”. In: *J Neurol* 257.10, pp. 1713–7. DOI: 10.1007/s00415-010-5609-1.
- Gousset, K. et al. (2009). “Prions hijack tunnelling nanotubes for intercellular spread”. In: *Nat Cell Biol* 11.3, pp. 328–36. DOI: 10.1038/ncb1841.
- Gowing, G., T. Philips, B. Van Wijmeersch, J. N. Audet, M. Dewil, L. Van Den Bosch, A. D. Billiau, W. Robberecht, and J. P. Julien (2008). “Ablation of proliferating microglia does not affect motor neuron degeneration in amyotrophic lateral sclerosis caused by mutant superoxide dismutase”. In: *J Neurosci* 28.41, pp. 10234–44. DOI: 10.1523/JNEUROSCI.3494-08.2008.
- Graaff, M. M. van der et al. (2011). “Upper and extra-motoneuron involvement in early motoneuron disease: a diffusion tensor imaging study”. In: *Brain* 134.Pt 4, pp. 1211–28. DOI: 10.1093/brain/awr016.
- Grad, L. I. et al. (2014). “Intercellular propagated misfolding of wild-type Cu/Zn superoxide dismutase occurs via exosome-dependent and -independent mechanisms”. In: *Proc Natl Acad Sci U S A* 111.9, pp. 3620–5. DOI: 10.1073/pnas.1312245111.
- Gravel, M., L. C. Beland, G. Soucy, E. Abdelhamid, R. Rahimian, C. Gravel, and J. Kriz (2016). “IL-10 Controls Early Microglial Phenotypes and Disease Onset in ALS Caused by Misfolded Superoxide Dismutase 1”. In: *J Neurosci* 36.3, pp. 1031–48. DOI: 10.1523/JNEUROSCI.0854-15.2016.
- Great Lakes A. L. S. Study Group, (2003). “A comparison of muscle strength testing techniques in amyotrophic lateral sclerosis”. In: *Neurology* 61.11, pp. 1503–7.
- Groen, E. J. et al. (2010). “FUS mutations in familial amyotrophic lateral sclerosis in the Netherlands”. In: *Arch Neurol* 67.2, pp. 224–30. DOI: 10.1001/archneurol.2009.329.

- Grossman, M. et al. (2014). “Phosphorylated tau as a candidate biomarker for amyotrophic lateral sclerosis”. In: *JAMA Neurol* 71.4, pp. 442–8. DOI: 10.1001/jamaneurol.2013.6064.
- Grossmann, J., B. Roschitzki, C. Panse, C. Fortes, S. Barkow-Oesterreicher, D. Rutishauser, and R. Schlapbach (2010). “Implementation and evaluation of relative and absolute quantification in shotgun proteomics with label-free methods”. In: *J Proteomics* 73.9, pp. 1740–6. DOI: 10.1016/j.jprot.2010.05.011.
- Gubbay, S. S., E. Kahana, N. Zilber, G. Cooper, S. Pintov, and Y. Leibowitz (1985). “Amyotrophic lateral sclerosis. A study of its presentation and prognosis”. In: *J Neurol* 232.5, pp. 295–300.
- Gui, Y., H. Liu, L. Zhang, W. Lv, and X. Hu (2015). “Altered microRNA profiles in cerebrospinal fluid exosome in Parkinson disease and Alzheimer disease”. In: *Oncotarget* 6.35, pp. 37043–53. DOI: 10.18632/oncotarget.6158.
- Guldbrandsen, A., H. Vethe, Y. Farag, E. Oveland, H. Garberg, M. Berle, K. M. Myhr, J. A. Opsahl, H. Barsnes, and F. S. Berven (2014). “In-depth characterization of the cerebrospinal fluid (CSF) proteome displayed through the CSF proteome resource (CSF-PR)”. In: *Mol Cell Proteomics* 13.11, pp. 3152–63. DOI: 10.1074/mcp.M114.038554.
- Gundry, R. L., M. Y. White, C. I. Murray, L. A. Kane, Q. Fu, B. A. Stanley, and J. E. Van Eyk (2009). “Preparation of proteins and peptides for mass spectrometry analysis in a bottom-up proteomics workflow”. In: *Curr Protoc Mol Biol* Chapter 10, Unit10 25. DOI: 10.1002/0471142727.mb1025s88.
- Gunnarsson, M. et al. (2011). “Axonal damage in relapsing multiple sclerosis is markedly reduced by natalizumab”. In: *Ann Neurol* 69.1, pp. 83–9. DOI: 10.1002/ana.22247.
- Gunther, R., E. Krause, M. Schumann, I. E. Blasig, and R. F. Haseloff (2015). “Depletion of highly abundant proteins from human cerebrospinal fluid: a cautionary note”. In: *Mol Neurodegener* 10, p. 53. DOI: 10.1186/s13024-015-0050-7.

- Guo, B. B., S. A. Bellingham, and A. F. Hill (2016). “Stimulating the Release of Exosomes Increases the Intercellular Transfer of Prions”. In: *J Biol Chem* 291.10, pp. 5128–37. DOI: 10.1074/jbc.M115.684258.
- Guo, J., X. Yang, L. Gao, and D. Zang (2017). “Evaluating the levels of CSF and serum factors in ALS”. In: *Brain Behav* 7.3, e00637. DOI: 10.1002/brb3.637.
- Hakansson, I., A. Tisell, P. Cassel, K. Blennow, H. Zetterberg, P. Lundberg, C. Dahle, M. Vrethem, and J. Ernerudh (2017). “Neurofilament light chain in cerebrospinal fluid and prediction of disease activity in clinically isolated syndrome and relapsing-remitting multiple sclerosis”. In: *Eur J Neurol* 24.5, pp. 703–712. DOI: 10.1111/ene.13274.
- Hara, T. et al. (2006). “Suppression of basal autophagy in neural cells causes neurodegenerative disease in mice”. In: *Nature* 441.7095, pp. 885–9. DOI: 10.1038/nature04724.
- Hawrylycz, M. J. et al. (2012). “An anatomically comprehensive atlas of the adult human brain transcriptome”. In: *Nature* 489.7416, pp. 391–399. DOI: 10.1038/nature11405.
- Hecht, M. J., F. Fellner, C. Fellner, M. J. Hilz, D. Heuss, and B. Neundorfer (2001). “MRI-FLAIR images of the head show corticospinal tract alterations in ALS patients more frequently than T2-, T1- and proton-density-weighted images”. In: *J Neurol Sci* 186.1-2, pp. 37–44.
- Heijnen, H. F., A. E. Schiel, R. Fijnheer, H. J. Geuze, and J. J. Sixma (1999). “Activated platelets release two types of membrane vesicles: microvesicles by surface shedding and exosomes derived from exocytosis of multivesicular bodies and alpha-granules”. In: *Blood* 94.11, pp. 3791–9.
- Highley, J. R. et al. (2014). “Loss of nuclear TDP-43 in amyotrophic lateral sclerosis (ALS) causes altered expression of splicing machinery and widespread dysregulation of RNA splicing in motor neurones”. In: *Neuropathol Appl Neurobiol* 40.6, pp. 670–85. DOI: 10.1111/nan.12148.

- Hinsinger, G. et al. (2015). “Chitinase 3-like proteins as diagnostic and prognostic biomarkers of multiple sclerosis”. In: *Mult Scler* 21.10, pp. 1251–61. DOI: 10.1177/1352458514561906.
- Hirano, M., Y. Nakamura, K. Saigoh, H. Sakamoto, S. Ueno, C. Isono, K. Miyamoto, M. Akamatsu, Y. Mitsui, and S. Kusunoki (2013). “Mutations in the gene encoding p62 in Japanese patients with amyotrophic lateral sclerosis”. In: *Neurology* 80.5, pp. 458–63. DOI: 10.1212/WNL.0b013e31827f0fe5.
- Ho, C. S., C. W. Lam, M. H. Chan, R. C. Cheung, L. K. Law, L. C. Lit, K. F. Ng, M. W. Suen, and H. L. Tai (2003). “Electrospray ionisation mass spectrometry: principles and clinical applications”. In: *Clin Biochem Rev* 24.1, pp. 3–12.
- Holm, S (1979). “A simple sequentially rejective multiple test procedure”. In: *Scandinavian Journal of Statistics* 6.2, pp. 65–70.
- Hondius, D. C., P. van Nierop, K. W. Li, J. J. Hoozemans, R. C. van der Schors, E. S. van Haastert, S. M. van der Vies, A. J. Rozemuller, and A. B. Smit (2016). “Profiling the human hippocampal proteome at all pathologic stages of Alzheimer’s disease”. In: *Alzheimers Dement* 12.6, pp. 654–68. DOI: 10.1016/j.jalz.2015.11.002.
- Hong, P., S. Koza, and E. S. Bouvier (2012). “Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and their Aggregates”. In: *J Liq Chromatogr Relat Technol* 35.20, pp. 2923–2950. DOI: 10.1080/10826076.2012.743724.
- Hu, Q., R. J. Noll, H. Li, A. Makarov, M. Hardman, and R. Graham Cooks (2005). “The Orbitrap: a new mass spectrometer”. In: *J Mass Spectrom* 40.4, pp. 430–43. DOI: 10.1002/jms.856.
- Huang, X. et al. (2013). “Characterization of human plasma-derived exosomal RNAs by deep sequencing”. In: *BMC Genomics* 14, p. 319. DOI: 10.1186/1471-2164-14-319.
- Hutchens TW; Yip, T-Y. (1993). “New Desorption Strategies for the Mass Spectrometric Analysis of Macromolecules”. In: *Rapid Commun Mass Spectrom* 7, pp. 576–580.

- Iguchi, Y. et al. (2016). “Exosome secretion is a key pathway for clearance of pathological TDP-43”. In: *Brain* 139.Pt 12, pp. 3187–3201. DOI: 10.1093/brain/aww237.
- Ince, P. G., J. Evans, M. Knopp, G. Forster, H. H. Hamdalla, S. B. Wharton, and P. J. Shaw (2003). “Corticospinal tract degeneration in the progressive muscular atrophy variant of ALS”. In: *Neurology* 60.8, pp. 1252–8.
- Irwin, D. J. et al. (2013). “Cognitive decline and reduced survival in C9orf72 expansion frontotemporal degeneration and amyotrophic lateral sclerosis”. In: *J Neurol Neurosurg Psychiatry* 84.2, pp. 163–9. DOI: 10.1136/jnnp-2012-303507.
- Ishii, A., R. Dutta, G. M. Wark, S. I. Hwang, D. K. Han, B. D. Trapp, S. E. Pfeiffer, and R. Bansal (2009). “Human myelin proteome and comparative analysis with mouse myelin”. In: *Proc Natl Acad Sci U S A* 106.34, pp. 14605–10. DOI: 10.1073/pnas.0905936106.
- Ishii, T., E. Kawakami, K. Endo, H. Misawa, and K. Watabe (2017). “Formation and spreading of TDP-43 aggregates in cultured neuronal and glial cells demonstrated by time-lapse imaging”. In: *PLoS One* 12.6, e0179375. DOI: 10.1371/journal.pone.0179375.
- Jacobsson, J., P. A. Jonsson, P. M. Andersen, L. Forsgren, and S. L. Marklund (2001). “Superoxide dismutase in CSF from amyotrophic lateral sclerosis patients with and without CuZn-superoxide dismutase mutations”. In: *Brain* 124.Pt 7, pp. 1461–6.
- Jakel, S. and L. Dimou (2017). “Glial Cells and Their Function in the Adult Brain: A Journey through the History of Their Ablation”. In: *Front Cell Neurosci* 11, p. 24. DOI: 10.3389/fncel.2017.00024.
- Jin, J., F. Hu, Q. Zhang, R. Jia, and J. Dang (2016). “Hyperintensity of the corticospinal tract on FLAIR: A simple and sensitive objective upper motor neuron degeneration marker in clinically verified amyotrophic lateral sclerosis”. In: *J Neurol Sci* 367, pp. 177–83. DOI: 10.1016/j.jns.2016.06.005.

- Johnson, B. S., D. Snead, J. J. Lee, J. M. McCaffery, J. Shorter, and A. D. Gitler (2009). “TDP-43 is intrinsically aggregation-prone, and amyotrophic lateral sclerosis-linked mutations accelerate aggregation and increase toxicity”. In: *J Biol Chem* 284.30, pp. 20329–39. DOI: 10.1074/jbc.M109.010264.
- Johnson, J. O. et al. (2010). “Exome sequencing reveals VCP mutations as a cause of familial ALS”. In: *Neuron* 68.5, pp. 857–64. DOI: 10.1016/j.neuron.2010.11.036.
- Johnson, J. O. et al. (2014). “Mutations in the Matrin 3 gene cause familial amyotrophic lateral sclerosis”. In: *Nat Neurosci* 17.5, pp. 664–666. DOI: 10.1038/nn.3688.
- Josephs, K. A. et al. (2017). “Rates of hippocampal atrophy and presence of post-mortem TDP-43 in patients with Alzheimer’s disease: a longitudinal retrospective study”. In: *Lancet Neurol* 16.11, pp. 917–924. DOI: 10.1016/S1474-4422(17)30284-3.
- Kabashi, E., J. N. Agar, M. J. Strong, and H. D. Durham (2012). “Impaired proteasome function in sporadic amyotrophic lateral sclerosis”. In: *Amyotroph Lateral Scler* 13.4, pp. 367–71. DOI: 10.3109/17482968.2012.686511.
- Kammers, K., R. N. Cole, C. Tiengwe, and I. Ruczinski (2015). “Detecting Significant Changes in Protein Abundance”. In: *EuPA Open Proteom* 7, pp. 11–19. DOI: 10.1016/j.euprot.2015.02.002.
- Kang, JS. (2012). “Principles and Applications of LC-MS/MS for the Quantitative Bioanalysis of Analytes in Various Biological Samples”. In: *Tandem Mass Spectrometry*. Ed. by J. Prasain. Vol. 1. InTechOpen. Chap. 20. DOI: 10.5772/32085.
- Kang, S. H., Y. Li, M. Fukaya, I. Lorenzini, D. W. Cleveland, L. W. Ostrow, J. D. Rothstein, and D. E. Bergles (2013). “Degeneration and impaired regeneration of gray matter oligodendrocytes in amyotrophic lateral sclerosis”. In: *Nat Neurosci* 16.5, pp. 571–9. DOI: 10.1038/nn.3357.

- Kanneganti, M., A. Kamba, and E. Mizoguchi (2012). “Role of chitotriosidase (chitinase 1) under normal and disease conditions”. In: *J Epithel Biol Pharmacol* 5, pp. 1–9.
- Kapp, E. and F. Schutz (2007). “Overview of tandem mass spectrometry (MS/MS) database search algorithms”. In: *Curr Protoc Protein Sci* Chapter 25, Unit25 2. DOI: 10.1002/0471140864.ps2502s49.
- Kasai, T., T. Tokuda, N. Ishigami, H. Sasayama, P. Foulds, D. J. Mitchell, D. M. Mann, D. Allsop, and M. Nakagawa (2009). “Increased TDP-43 protein in cerebrospinal fluid of patients with amyotrophic lateral sclerosis”. In: *Acta Neuropathol* 117.1, pp. 55–62. DOI: 10.1007/s00401-008-0456-1.
- Kaufmann, P., G. Levy, J. L. Thompson, M. L. Delbene, V. Battista, P. H. Gordon, L. P. Rowland, B. Levin, and H. Mitsumoto (2005). “The ALSFRS_r predicts survival time in an ALS clinic population”. In: *Neurology* 64.1, pp. 38–43. DOI: 10.1212/01.WNL.0000148648.38313.64.
- Kawamata, T., H. Akiyama, T. Yamada, and P. L. McGeer (1992). “Immunologic reactions in amyotrophic lateral sclerosis brain and spinal cord tissue”. In: *Am J Pathol* 140.3, pp. 691–707.
- Keilhauer, E. C., M. Y. Hein, and M. Mann (2015). “Accurate protein complex retrieval by affinity enrichment mass spectrometry (AE-MS) rather than affinity purification mass spectrometry (AP-MS)”. In: *Mol Cell Proteomics* 14.1, pp. 120–35. DOI: 10.1074/mcp.M114.041012.
- Keller, J. N., K. B. Hanni, and W. R. Markesbery (2000). “Possible involvement of proteasome inhibition in aging: implications for oxidative stress”. In: *Mech Ageing Dev* 113.1, pp. 61–70.
- Kenna, K. P. et al. (2016). “NEK1 variants confer susceptibility to amyotrophic lateral sclerosis”. In: *Nat Genet* 48.9, pp. 1037–42. DOI: 10.1038/ng.3626.
- Keshishian, H., M. W. Burgess, M. A. Gillette, P. Mertins, K. R. Clauser, D. R. Mani, E. W. Kuhn, L. A. Farrell, R. E. Gerszten, and S. A. Carr (2015). “Multiplexed, Quantitative Workflow for Sensitive Biomarker Discovery in Plasma Yields

- Novel Candidates for Early Myocardial Injury”. In: *Mol Cell Proteomics* 14.9, pp. 2375–93. DOI: 10.1074/mcp.M114.046813.
- Kida, E., S. Palminiello, A. A. Golabek, M. Walus, T. Wierzba-Bobrowicz, A. Rabe, G. Albertini, and K. E. Wisniewski (2006). “Carbonic anhydrase II in the developing and adult human brain”. In: *J Neuropathol Exp Neurol* 65.7, pp. 664–74. DOI: 10.1097/01.jnen.0000225905.52002.3e.
- Kiernan, M. C., S. Vucic, B. C. Cheah, M. R. Turner, A. Eisen, O. Hardiman, J. R. Burrell, and M. C. Zoing (2011). “Amyotrophic lateral sclerosis”. In: *Lancet* 377.9769, pp. 942–55. DOI: 10.1016/S0140-6736(10)61156-7.
- Kim, H. J. et al. (2013). “Mutations in prion-like domains in hnRNPA2B1 and hnRNPA1 cause multisystem proteinopathy and ALS”. In: *Nature* 495.7442, pp. 467–73. DOI: 10.1038/nature11922.
- Kim, J., J. M. Basak, and D. M. Holtzman (2009). “The role of apolipoprotein E in Alzheimer’s disease”. In: *Neuron* 63.3, pp. 287–303. DOI: 10.1016/j.neuron.2009.06.026.
- King, O. D., A. D. Gitler, and J. Shorter (2012). “The tip of the iceberg: RNA-binding proteins with prion-like domains in neurodegenerative disease”. In: *Brain Res* 1462, pp. 61–80. DOI: 10.1016/j.brainres.2012.01.016.
- Kirby, J. et al. (2011). “Phosphatase and tensin homologue/protein kinase B pathway linked to motor neuron survival in human superoxide dismutase 1-related amyotrophic lateral sclerosis”. In: *Brain* 134.Pt 2, pp. 506–17. DOI: 10.1093/brain/awq345.
- Knibb, J. A., N. Keren, A. Kulka, P. N. Leigh, S. Martin, C. E. Shaw, M. Tsuda, and A. Al-Chalabi (2016). “A clinical tool for predicting survival in ALS”. In: *J Neurol Neurosurg Psychiatry* 87.12, pp. 1361–1367. DOI: 10.1136/jnnp-2015-312908.
- Kordower, J. H., Y. Chu, R. A. Hauser, T. B. Freeman, and C. W. Olanow (2008). “Lewy body-like pathology in long-term embryonic nigral transplants in Parkinson’s disease”. In: *Nat Med* 14.5, pp. 504–6. DOI: 10.1038/nm1747.

- Kowal, J., G. Arras, M. Colombo, M. Jouve, J. P. Morath, B. Primdal-Bengtson, F. Dingli, D. Loew, M. Tkach, and C. Thery (2016). “Proteomic comparison defines novel markers to characterize heterogeneous populations of extracellular vesicle subtypes”. In: *Proc Natl Acad Sci U S A* 113.8, E968–77. DOI: 10.1073/pnas.1521230113.
- Kramer-Albers, E. M., N. Bretz, S. Tenzer, C. Winterstein, W. Mobius, H. Berger, K. A. Nave, H. Schild, and J. Trotter (2007). “Oligodendrocytes secrete exosomes containing major myelin and stress-protective proteins: Trophic support for axons?” In: *Proteomics Clin Appl* 1.11, pp. 1446–61. DOI: 10.1002/prca.200700522.
- Krey, J. F., P. A. Wilmarth, J. B. Shin, J. Klimek, N. E. Sherman, E. D. Jeffery, D. Choi, L. L. David, and P. G. Barr-Gillespie (2014). “Accurate label-free protein quantitation with high- and low-resolution mass spectrometers”. In: *J Proteome Res* 13.2, pp. 1034–1044. DOI: 10.1021/pr401017h.
- Kuffner, R. et al. (2015). “Crowdsourced analysis of clinical trial data to predict amyotrophic lateral sclerosis progression”. In: *Nat Biotechnol* 33.1, pp. 51–7. DOI: 10.1038/nbt.3051.
- Kuhn, E., J. Wu, J. Karl, H. Liao, W. Zolg, and B. Guild (2004). “Quantification of C-reactive protein in the serum of patients with rheumatoid arthritis using multiple reaction monitoring mass spectrometry and ^{13}C -labeled peptide standards”. In: *Proteomics* 4.4, pp. 1175–86. DOI: 10.1002/pmic.200300670.
- Kunz, L. I., E. F. van’t Wout, A. van Schadewijk, D. S. Postma, H. A. Kerstjens, P. J. Sterk, and P. S. Hiemstra (2015). “Regulation of YKL-40 expression by corticosteroids: effect on pro-inflammatory macrophages in vitro and its modulation in COPD in vivo”. In: *Respir Res* 16, p. 154. DOI: 10.1186/s12931-015-0314-3.
- Kushner, P. D., D. T. Stephenson, and S. Wright (1991). “Reactive astrogliosis is widespread in the subcortical white matter of amyotrophic lateral sclerosis brain”. In: *J Neuropathol Exp Neurol* 50.3, pp. 263–77.

- Kwiatkowski T. J., Jr. et al. (2009). “Mutations in the FUS/TLS gene on chromosome 16 cause familial amyotrophic lateral sclerosis”. In: *Science* 323.5918, pp. 1205–8. DOI: 10.1126/science.1166066.
- Labra, J., P. Menon, K. Byth, S. Morrison, and S. Vucic (2016). “Rate of disease progression: a prognostic biomarker in ALS”. In: *J Neurol Neurosurg Psychiatry* 87.6, pp. 628–32. DOI: 10.1136/jnnp-2015-310998.
- Lagier-Tourenne, C., M. Polymenidou, and D. W. Cleveland (2010). “TDP-43 and FUS/TLS: emerging roles in RNA processing and neurodegeneration”. In: *Hum Mol Genet* 19.R1, R46–64. DOI: 10.1093/hmg/ddq137.
- Lagier-Tourenne, C. et al. (2012). “Divergent roles of ALS-linked proteins FUS/TLS and TDP-43 intersect in processing long pre-mRNAs”. In: *Nat Neurosci* 15.11, pp. 1488–97. DOI: 10.1038/nn.3230.
- Lagier-Tourenne, C. et al. (2013). “Targeted degradation of sense and antisense C9orf72 RNA foci as therapy for ALS and frontotemporal degeneration”. In: *Proc Natl Acad Sci U S A* 110.47, E4530–9. DOI: 10.1073/pnas.1318835110.
- Lange, V., P. Picotti, B. Domon, and R. Aebersold (2008). “Selected reaction monitoring for quantitative proteomics: a tutorial”. In: *Mol Syst Biol* 4, p. 222. DOI: 10.1038/msb.2008.61.
- Lazar, C., L. Gatto, M. Ferro, C. Bruley, and T. Burger (2016). “Accounting for the Multiple Natures of Missing Values in Label-Free Quantitative Proteomics Data Sets to Compare Imputation Strategies”. In: *J Proteome Res* 15.4, pp. 1116–25. DOI: 10.1021/acs.jproteome.5b00981.
- Le Ber, I. et al. (2013). “SQSTM1 mutations in French patients with frontotemporal dementia or frontotemporal dementia with amyotrophic lateral sclerosis”. In: *JAMA Neurol* 70.11, pp. 1403–10. DOI: 10.1001/jamaneurol.2013.3849.
- Le Forestier, N., T. Maisonobe, A. Piquard, S. Rivaud, L. Crevier-Buchman, F. Salachas, P. F. Pradat, L. Lacomblez, and V. Meininger (2001). “Does primary lateral sclerosis exist? A study of 20 patients and a review of the literature”. In: *Brain* 124.Pt 10, pp. 1989–99.

- Leblond, C. S. et al. (2016). “Replication study of MATR3 in familial and sporadic amyotrophic lateral sclerosis”. In: *Neurobiol Aging* 37, 209 e17-209 e21. DOI: 10.1016/j.neurobiolaging.2015.09.013.
- Lee, Y. et al. (2012). “Oligodendroglia metabolically support axons and contribute to neurodegeneration”. In: *Nature* 487.7408, pp. 443–8. DOI: 10.1038/nature11314.
- Lefterov, I. M., R. P. Koldamova, and J. S. Lazo (2000). “Human bleomycin hydrolase regulates the secretion of amyloid precursor protein”. In: *FASEB J* 14.12, pp. 1837–47.
- Legha, S. S., H. L. Davis, and F. M. Muggia (1978). “Hormonal therapy of breast cancer: new approaches and concepts”. In: *Ann Intern Med* 88.1, pp. 69–77.
- Leonardi, A., G. Abbruzzese, L. Arata, L. Cocito, and M. Vische (1984). “Cerebrospinal fluid (CSF) findings in amyotrophic lateral sclerosis”. In: *J Neurol* 231.2, pp. 75–8.
- Levin, Y. (2011). “The role of statistical power analysis in quantitative proteomics”. In: *Proteomics* 11.12, pp. 2565–7. DOI: 10.1002/pmic.201100033.
- Lewczuk, P. et al. (2018). “Cerebrospinal fluid and blood biomarkers for neurodegenerative dementias: An update of the Consensus of the Task Force on Biological Markers in Psychiatry of the World Federation of Societies of Biological Psychiatry”. In: *World J Biol Psychiatry* 19.4, pp. 244–328. DOI: 10.1080/15622975.2017.1375556.
- Li, Q. and B. A. Barres (2018). “Microglia and macrophages in brain homeostasis and disease”. In: *Nat Rev Immunol* 18.4, pp. 225–242. DOI: 10.1038/nri.2017.125.
- Li, Z., R. M. Adams, K. Chourey, G. B. Hurst, R. L. Hettich, and C. Pan (2012). “Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos”. In: *J Proteome Res* 11.3, pp. 1582–90. DOI: 10.1021/pr200748h.
- Lin, D., W. E. Alborn, R. J. Slebos, and D. C. Liebler (2013). “Comparison of protein immunoprecipitation-multiple reaction monitoring with ELISA for assay

- of biomarker candidates in plasma”. In: *J Proteome Res* 12.12, pp. 5996–6003. DOI: 10.1021/pr400877e.
- Lipinski, M. M. et al. (2010). “Genome-wide analysis reveals mechanisms modulating autophagy in normal brain aging and in Alzheimer’s disease”. In: *Proc Natl Acad Sci U S A* 107.32, pp. 14164–9. DOI: 10.1073/pnas.1009485107.
- Liu, C. (2011). “The application of SELDI-TOF-MS in clinical diagnosis of cancers”. In: *J Biomed Biotechnol* 2011, p. 245821. DOI: 10.1155/2011/245821.
- Liu, T. et al. (2018). “Tumor-associated macrophages in human breast cancer produce new monocyte attracting and pro-angiogenic factor YKL-39 indicative for increased metastasis after neoadjuvant chemotherapy”. In: *Oncoimmunology* 7.6, e1436922. DOI: 10.1080/2162402X.2018.1436922.
- Liu, Y., A. Beyer, and R. Aebersold (2016). “On the Dependency of Cellular Protein Levels on mRNA Abundance”. In: *Cell* 165.3, pp. 535–50. DOI: 10.1016/j.cell.2016.03.014.
- Liu-Yesucevitz, L. et al. (2010). “Tar DNA binding protein-43 (TDP-43) associates with stress granules: analysis of cultured cells and pathological brain tissue”. In: *PLoS One* 5.10, e13250. DOI: 10.1371/journal.pone.0013250.
- Lobb, R. J., M. Becker, S. W. Wen, C. S. Wong, A. P. Wiegman, A. Leimgruber, and A. Moller (2015). “Optimized exosome isolation protocol for cell culture supernatant and human plasma”. In: *J Extracell Vesicles* 4, p. 27031. DOI: 10.3402/jev.v4.27031.
- Logroscino, G., E. Beghi, S. Zoccolella, R. Palagano, A. Fraddosio, I. L. Simone, P. Lamberti, V. Lepore, L. Serlenga, and Slap Registry (2005). “Incidence of amyotrophic lateral sclerosis in southern Italy: a population based study”. In: *J Neurol Neurosurg Psychiatry* 76.8, pp. 1094–8. DOI: 10.1136/jnnp.2004.039180.
- Logroscino, G., B. J. Traynor, O. Hardiman, A. Chio, D. Mitchell, R. J. Swingler, A. Millul, E. Benn, E. Beghi, and Eurals (2010). “Incidence of amyotrophic lateral sclerosis in Europe”. In: *J Neurol Neurosurg Psychiatry* 81.4, pp. 385–90. DOI: 10.1136/jnnp.2009.183525.

- Lowe, J. (1994). “New pathological findings in amyotrophic lateral sclerosis”. In: *J Neurol Sci* 124 Suppl, pp. 38–51.
- Lu, C. H. et al. (2015). “Neurofilament light chain: A prognostic biomarker in amyotrophic lateral sclerosis”. In: *Neurology* 84.22, pp. 2247–57. DOI: 10.1212/WNL.0000000000001642.
- Maetzler, W., J. Michelis, J. Tomiuk, A. Melms, C. Becker, T. Gasser, C. Schulte, and D. Berg (2009). “A single-nucleotide polymorphism of the osteopontin gene may contribute to a susceptibility to Lewy body disease”. In: *J Neural Transm (Vienna)* 116.5, pp. 599–605. DOI: 10.1007/s00702-009-0209-x.
- Majounie, E. et al. (2012). “Frequency of the C9orf72 hexanucleotide repeat expansion in patients with amyotrophic lateral sclerosis and frontotemporal dementia: a cross-sectional study”. In: *Lancet Neurol* 11.4, pp. 323–30. DOI: 10.1016/S1474-4422(12)70043-1.
- Malessa, S., P. N. Leigh, O. Bertel, E. Sluga, and O. Hornykiewicz (1991). “Amyotrophic lateral sclerosis: glutamate dehydrogenase and transmitter amino acids in the spinal cord”. In: *J Neurol Neurosurg Psychiatry* 54.11, pp. 984–8.
- March, R. E. (2009). “Quadrupole ion traps”. In: *Mass Spectrom Rev* 28.6, pp. 961–89. DOI: 10.1002/mas.20250.
- Marengo, E. and E. Robotti (2014). “Biomarkers for pancreatic cancer: recent achievements in proteomics and genomics through classical and multivariate statistical methods”. In: *World J Gastroenterol* 20.37, pp. 13325–42. DOI: 10.3748/wjg.v20.i37.13325.
- Mariosa, D., N. Hammar, H. Malmstrom, C. Ingre, I. Jungner, W. Ye, F. Fang, and G. Walldius (2017). “Blood biomarkers of carbohydrate, lipid, and apolipoprotein metabolisms and risk of amyotrophic lateral sclerosis: A more than 20-year follow-up of the Swedish AMORIS cohort”. In: *Ann Neurol* 81.5, pp. 718–728. DOI: 10.1002/ana.24936.
- Martin, J. H. (2005). “The corticospinal system: from development to motor control”. In: *Neuroscientist* 11.2, pp. 161–73. DOI: 10.1177/1073858404270843.

- Martinez-Muriana, A., R. Mancuso, I. Francos-Quijorna, A. Olmos-Alonso, R. Osta, V. H. Perry, X. Navarro, D. Gomez-Nicola, and R. Lopez-Vales (2016). “CSF1R blockade slows the progression of amyotrophic lateral sclerosis by reducing microgliosis and invasion of macrophages into peripheral nerves”. In: *Sci Rep* 6, p. 25663. DOI: 10.1038/srep25663.
- Maruyama, H. et al. (2010). “Mutations of optineurin in amyotrophic lateral sclerosis”. In: *Nature* 465.7295, pp. 223–6. DOI: 10.1038/nature08971.
- McCombe, P. A., C. Pfluger, P. Singh, C. Y. Lim, C. Airey, and R. D. Henderson (2015). “Serial measurements of phosphorylated neurofilament-heavy in the serum of subjects with amyotrophic lateral sclerosis”. In: *J Neurol Sci* 353.1-2, pp. 122–9. DOI: 10.1016/j.jns.2015.04.032.
- McGeer, P. L., T. Kawamata, D. G. Walker, H. Akiyama, I. Tooyama, and E. G. McGeer (1993). “Microglia in degenerative neurological disease”. In: *Glia* 7.1, pp. 84–92. DOI: 10.1002/glia.440070114.
- McKee, A. C., T. D. Stein, P. T. Kiernan, and V. E. Alvarez (2015). “The neuropathology of chronic traumatic encephalopathy”. In: *Brain Pathol* 25.3, pp. 350–64. DOI: 10.1111/bpa.12248.
- Megger, D. A., T. Bracht, H. E. Meyer, and B. Sitek (2013). “Label-free quantification in clinical proteomics”. In: *Biochim Biophys Acta* 1834.8, pp. 1581–90. DOI: 10.1016/j.bbapap.2013.04.001.
- Mendonca, D. M., L. Chimelli, and A. M. Martinez (2006). “Expression of ubiquitin and proteasome in motoneurons and astrocytes of spinal cords from patients with amyotrophic lateral sclerosis”. In: *Neurosci Lett* 404.3, pp. 315–9. DOI: 10.1016/j.neulet.2006.06.009.
- Mendonca, D. M., L. Pizzati, K. Mostacada, S. Martins S. C. de, R. Higashi, L. Ayres Sa, V. Moura Neto, L. Chimelli, and A. M. Martinez (2012). “Neuroproteomics: an insight into ALS”. In: *Neurol Res* 34.10, pp. 937–43. DOI: 10.1179/1743132812Y.0000000092.

- Menke, R. A., E. Gray, C. H. Lu, J. Kuhle, K. Talbot, A. Malaspina, and M. R. Turner (2015). “CSF neurofilament light chain reflects corticospinal tract degeneration in ALS”. In: *Ann Clin Transl Neurol* 2.7, pp. 748–55. DOI: 10.1002/acn3.212.
- Merner, N. D. et al. (2012). “Exome sequencing identifies FUS mutations as a cause of essential tremor”. In: *Am J Hum Genet* 91.2, pp. 313–9. DOI: 10.1016/j.ajhg.2012.07.002.
- Miller, T. M. et al. (2013). “An antisense oligonucleotide against SOD1 delivered intrathecally for patients with SOD1 familial amyotrophic lateral sclerosis: a phase 1, randomised, first-in-man study”. In: *Lancet Neurol* 12.5, pp. 435–42. DOI: 10.1016/S1474-4422(13)70061-9.
- Mitchell, R. M., W. M. Freeman, W. T. Randazzo, H. E. Stephens, J. L. Beard, Z. Simmons, and J. R. Connor (2009). “A CSF biomarker panel for identification of patients with amyotrophic lateral sclerosis”. In: *Neurology* 72.1, pp. 14–9. DOI: 10.1212/01.wnl.0000333251.36681.a5.
- Mitsumoto, H. et al. (2015). “Phenotypic and molecular analyses of primary lateral sclerosis”. In: *Neurol Genet* 1.1, e3. DOI: 10.1212/01.NXG.0000464294.88607.dd.
- Mizuno, Y., M. Amari, M. Takatama, H. Aizawa, B. Mihara, and K. Okamoto (2006). “Immunoreactivities of p62, an ubiquitin-binding protein, in the spinal anterior horn cells of patients with amyotrophic lateral sclerosis”. In: *J Neurol Sci* 249.1, pp. 13–8. DOI: 10.1016/j.jns.2006.05.060.
- Molliex, A., J. Temirov, J. Lee, M. Coughlin, A. P. Kanagaraj, H. J. Kim, T. Mittag, and J. P. Taylor (2015). “Phase separation by low complexity domains promotes stress granule assembly and drives pathological fibrillization”. In: *Cell* 163.1, pp. 123–33. DOI: 10.1016/j.cell.2015.09.015.
- Moore D. H., 2nd and R. G. Miller (2004). “Improving efficiency of ALS clinical trials using lead-in designs”. In: *Amyotroph Lateral Scler Other Motor Neuron Disord* 5 Suppl 1, pp. 57–60. DOI: 10.1080/17434470410019997.

- Mori, K. et al. (2013). “The C9orf72 GGGGCC repeat is translated into aggregating dipeptide-repeat proteins in FTL D/ALS”. In: *Science* 339.6125, pp. 1335–8. DOI: 10.1126/science.1232927.
- Motakis, E. S., G. P. Nason, P. Fryzlewicz, and G. A. Rutter (2006). “Variance stabilization and normalization for one-color microarray data using a data-driven multiscale approach”. In: *Bioinformatics* 22.20, pp. 2547–53. DOI: 10.1093/bioinformatics/btl412.
- Munch, C. et al. (2005). “Heterozygous R1101K mutation of the DCTN1 gene in a family with ALS and FTD”. In: *Ann Neurol* 58.5, pp. 777–80. DOI: 10.1002/ana.20631.
- Munch, C. et al. (2004). “Point mutations of the p150 subunit of dynactin (DCTN1) gene in ALS”. In: *Neurology* 63.4, pp. 724–6.
- Munsat, T. L., P. L. Andres, L. Finison, T. Conlon, and L. Thibodeau (1988). “The natural history of motoneuron loss in amyotrophic lateral sclerosis”. In: *Neurology* 38.3, pp. 409–13.
- Muralidharan-Chari, V., J. Clancy, C. Plou, M. Romao, P. Chavrier, G. Raposo, and C. D’Souza-Schorey (2009). “ARF6-regulated shedding of tumor cell-derived plasma membrane microvesicles”. In: *Curr Biol* 19.22, pp. 1875–85. DOI: 10.1016/j.cub.2009.09.059.
- Murray, M. E. et al. (2011). “Clinical and neuropathologic heterogeneity of c9FTD/ALS associated with hexanucleotide repeat expansion in C9ORF72”. In: *Acta Neuropathol* 122.6, pp. 673–90. DOI: 10.1007/s00401-011-0907-y.
- Mustapic, M., E. Eitan, Jr. Werner J. K., S. T. Berkowitz, M. P. Lazaropoulos, J. Tran, E. J. Goetzl, and D. Kapogiannis (2017). “Plasma Extracellular Vesicles Enriched for Neuronal Origin: A Potential Window into Brain Pathologic Processes”. In: *Front Neurosci* 11, p. 278. DOI: 10.3389/fnins.2017.00278.
- Nagy, D., T. Kato, and P. D. Kushner (1994). “Reactive astrocytes are widespread in the cortical gray matter of amyotrophic lateral sclerosis”. In: *J Neurosci Res* 38.3, pp. 336–47. DOI: 10.1002/jnr.490380312.

- Nesvizhskii, A. I. and R. Aebersold (2005). “Interpretation of shotgun proteomic data: the protein inference problem”. In: *Mol Cell Proteomics* 4.10, pp. 1419–40. DOI: 10.1074/mcp.R500012-MCP200.
- Neuhoff, N. von et al. (2012). “Monitoring CSF proteome alterations in amyotrophic lateral sclerosis: obstacles and perspectives in translating a novel marker panel to the clinic”. In: *PLoS One* 7.9, e44401. DOI: 10.1371/journal.pone.0044401.
- Neumann, M. et al. (2006). “Ubiquitinated TDP-43 in frontotemporal lobar degeneration and amyotrophic lateral sclerosis”. In: *Science* 314.5796, pp. 130–3. DOI: 10.1126/science.1134108.
- Niel, G. van et al. (2015). “Apolipoprotein E Regulates Amyloid Formation within Endosomes of Pigment Cells”. In: *Cell Rep* 13.1, pp. 43–51. DOI: 10.1016/j.celrep.2015.08.057.
- Nonaka, T. et al. (2013). “Prion-like properties of pathological TDP-43 aggregates from diseased brains”. In: *Cell Rep* 4.1, pp. 124–34. DOI: 10.1016/j.celrep.2013.06.007.
- Nordin, J. Z. et al. (2015). “Ultrafiltration with size-exclusion liquid chromatography for high yield isolation of extracellular vesicles preserving intact biophysical and functional properties”. In: *Nanomedicine* 11.4, pp. 879–83. DOI: 10.1016/j.nano.2015.01.003.
- Norris, F., R. Shepherd, E. Denys, K. U, E. Mukai, L. Elias, D. Holden, and H. Norris (1993). “Onset, natural history and outcome in idiopathic adult motor neuron disease”. In: *J Neurol Sci* 118.1, pp. 48–55.
- Noto, Y. et al. (2011). “Elevated CSF TDP-43 levels in amyotrophic lateral sclerosis: specificity, sensitivity, and a possible prognostic value”. In: *Amyotroph Lateral Scler* 12.2, pp. 140–3. DOI: 10.3109/17482968.2010.541263.
- Oberg, A. L. and O. Vitek (2009). “Statistical design of quantitative mass spectrometry-based proteomic experiments”. In: *J Proteome Res* 8.5, pp. 2144–56. DOI: 10.1021/pr8010099.

- Oeckl, P. et al. (2018). “Different neuroinflammatory profile in amyotrophic lateral sclerosis and frontotemporal dementia is linked to the clinical phase”. In: *J Neurol Neurosurg Psychiatry*. DOI: 10.1136/jnnp-2018-318868.
- O’Farrell, P. A., F. Gonzalez, W. Zheng, S. A. Johnston, and L. Joshua-Tor (1999). “Crystal structure of human bleomycin hydrolase, a self-compartmentalizing cysteine protease”. In: *Structure* 7.6, pp. 619–27.
- Ohgomori, T., J. Yamada, H. Takeuchi, K. Kadomatsu, and S. Jinno (2016). “Comparative morphometric analysis of microglia in the spinal cord of SOD1(G93A) transgenic mouse model of amyotrophic lateral sclerosis”. In: *Eur J Neurosci* 43.10, pp. 1340–51. DOI: 10.1111/ejn.13227.
- Ojha, J., G. Masilamoni, D. Dunlap, R. A. Udoff, and A. G. Cashikar (2011). “Sequestration of toxic oligomers by HspB1 as a cytoprotective mechanism”. In: *Mol Cell Biol* 31.15, pp. 3146–57. DOI: 10.1128/MCB.01187-10.
- Olsen, J. V., B. Macek, O. Lange, A. Makarov, S. Horning, and M. Mann (2007). “Higher-energy C-trap dissociation for peptide modification analysis”. In: *Nat Methods* 4.9, pp. 709–12. DOI: 10.1038/nmeth1060.
- Order, S. E. (1977). “Beneficial and detrimental effects of therapy on immunity in breast cancer”. In: *Int J Radiat Oncol Biol Phys* 2.3-4, pp. 377–80.
- O’Rourke, J. G. et al. (2016). “C9orf72 is required for proper macrophage and microglial function in mice”. In: *Science* 351.6279, pp. 1324–9. DOI: 10.1126/science.aaf1064.
- Parker, C. E. and C. H. Borchers (2014). “Mass spectrometry based biomarker discovery, verification, and validation—quality assurance and control of protein biomarker assays”. In: *Mol Oncol* 8.4, pp. 840–58. DOI: 10.1016/j.molonc.2014.03.006.
- Parkkila, S., A. K. Parkkila, H. Rajaniemi, G. N. Shah, J. H. Grubb, A. Waheed, and W. S. Sly (2001). “Expression of membrane-associated carbonic anhydrase XIV on neurons and axons in mouse and human brain”. In: *Proc Natl Acad Sci U S A* 98.4, pp. 1918–23. DOI: 10.1073/pnas.98.4.1918.

- Pasinetti, G. M. et al. (2006). “Identification of potential CSF biomarkers in ALS”. In: *Neurology* 66.8, pp. 1218–22. DOI: 10.1212/01.wnl.0000203129.82104.07.
- Patel, B. B., C. A. Barrero, A. Braverman, P. D. Kim, K. A. Jones, D. E. Chen, R. P. Bowler, S. Merali, S. G. Kelsen, and A. T. Yeung (2012). “Assessment of two immunodepletion methods: off-target effects and variations in immunodepletion efficiency may confound plasma proteomics”. In: *J Proteome Res* 11.12, pp. 5947–58. DOI: 10.1021/pr300686k.
- Phukan, J., M. Elamin, P. Bede, N. Jordan, L. Gallagher, S. Byrne, C. Lynch, N. Pender, and O. Hardiman (2012). “The syndrome of cognitive impairment in amyotrophic lateral sclerosis: a population-based study”. In: *J Neurol Neurosurg Psychiatry* 83.1, pp. 102–8. DOI: 10.1136/jnnp-2011-300188.
- Poesen, K. et al. (2017). “Neurofilament markers for ALS correlate with extent of upper and lower motor neuron disease”. In: *Neurology* 88.24, pp. 2302–2309. DOI: 10.1212/WNL.0000000000004029.
- Polymenidou, M. et al. (2011). “Long pre-mRNA depletion and RNA missplicing contribute to neuronal vulnerability from loss of TDP-43”. In: *Nat Neurosci* 14.4, pp. 459–68. DOI: 10.1038/nn.2779.
- Porter, K. A. (1957). “Effect of homologous bone marrow injections in x-irradiated rabbits”. In: *Br J Exp Pathol* 38.4, pp. 401–12.
- Pratt, J. M., D. M. Simpson, M. K. Doherty, J. Rivers, S. J. Gaskell, and R. J. Beynon (2006). “Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes”. In: *Nat Protoc* 1.2, pp. 1029–43. DOI: 10.1038/nprot.2006.129.
- Pringle, C. E., A. J. Hudson, D. G. Munoz, J. A. Kiernan, W. F. Brown, and G. C. Ebers (1992). “Primary lateral sclerosis. Clinical features, neuropathology and diagnostic criteria”. In: *Brain* 115 (Pt 2), pp. 495–520.
- Proudfoot, M., A. Jones, K. Talbot, A. Al-Chalabi, and M. R. Turner (2016). “The ALSFRS as an outcome measure in therapeutic trials and its relationship to symp-

- tom onset”. In: *Amyotroph Lateral Scler Frontotemporal Degener* 17.5-6, pp. 414–25. DOI: 10.3109/21678421.2016.1140786.
- Prudencio, M., P. J. Hart, D. R. Borchelt, and P. M. Andersen (2009). “Variation in aggregation propensities among ALS-associated variants of SOD1: correlation to human disease”. In: *Hum Mol Genet* 18.17, pp. 3217–26. DOI: 10.1093/hmg/ddp260.
- Prudencio, M. et al. (2015). “Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS”. In: *Nat Neurosci* 18.8, pp. 1175–82. DOI: 10.1038/nn.4065.
- Prusiner, S. B. (1982). “Novel proteinaceous infectious particles cause scrapie”. In: *Science* 216.4542, pp. 136–44.
- Puls, I. et al. (2003). “Mutant dynactin in motor neuron disease”. In: *Nat Genet* 33.4, pp. 455–6. DOI: 10.1038/ng1123.
- Querol-Vilaseca, M., M. Colom-Cadena, J. Pegueroles, C. San Martin-Paniello, J. Clarimon, O. Belbin, J. Fortea, and A. Lleó (2017). “YKL-40 (Chitinase 3-like I) is expressed in a subset of astrocytes in Alzheimer’s disease and other tauopathies”. In: *J Neuroinflammation* 14.1, p. 118. DOI: 10.1186/s12974-017-0893-7.
- Ramstrom, M., I. Ivonin, A. Johansson, H. Askmark, K. E. Markides, R. Zubarev, P. Hakansson, S. M. Aquilonius, and J. Bergquist (2004). “Cerebrospinal fluid protein patterns in neurodegenerative disease revealed by liquid chromatography-Fourier transform ion cyclotron resonance mass spectrometry”. In: *Proteomics* 4.12, pp. 4010–8. DOI: 10.1002/pmic.200400871.
- Ranganathan, S., G. C. Nicholl, S. Henry, F. Lutka, R. Sathanoori, D. Lacomis, and R. Bowser (2007). “Comparative proteomic profiling of cerebrospinal fluid between living and post mortem ALS and control subjects”. In: *Amyotroph Lateral Scler* 8.6, pp. 373–9. DOI: 10.1080/17482960701549681.
- Ranganathan, S., E. Williams, P. Ganchev, V. Gopalakrishnan, D. Lacomis, L. Urbinelli, K. Newhall, M. E. Cudkowicz, Jr. Brown R. H., and R. Bowser (2005). “Proteomic profiling of cerebrospinal fluid identifies biomarkers for amyotrophic

- lateral sclerosis". In: *J Neurochem* 95.5, pp. 1461–71. DOI: 10.1111/j.1471-4159.2005.03478.x.
- Ratovitski, T., E. Chighladze, E. Waldron, R. R. Hirschhorn, and C. A. Ross (2011). "Cysteine proteases bleomycin hydrolase and cathepsin Z mediate N-terminal proteolysis and toxicity of mutant huntingtin". In: *J Biol Chem* 286.14, pp. 12578–89. DOI: 10.1074/jbc.M110.185348.
- Rauniyar, N. and 3rd Yates J. R. (2014). "Isobaric labeling-based relative quantification in shotgun proteomics". In: *J Proteome Res* 13.12, pp. 5293–309. DOI: 10.1021/pr500880b.
- Ravits, J. M. and A. R. La Spada (2009). "ALS motor phenotype heterogeneity, focality, and spread: deconstructing motor neuron degeneration". In: *Neurology* 73.10, pp. 805–11. DOI: 10.1212/WNL.0b013e3181b6bbbd.
- Ravits, J., P. Paul, and C. Jorg (2007). "Focality of upper and lower motor neuron degeneration at the clinical onset of ALS". In: *Neurology* 68.19, pp. 1571–5. DOI: 10.1212/01.wnl.0000260965.20021.47.
- Reiber, H. (1980). "The discrimination between different blood-CSF barrier dysfunctions and inflammatory reactions of the CNS by a recent evaluation graph for the protein profile of cerebrospinal fluid". In: *J Neurol* 224.2, pp. 89–99.
- Renton, A. E. et al. (2011). "A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD". In: *Neuron* 72.2, pp. 257–68. DOI: 10.1016/j.neuron.2011.09.010.
- Rentzos, M. et al. (2010). "Interleukin-17 and interleukin-23 are elevated in serum and cerebrospinal fluid of patients with ALS: a reflection of Th17 cells activation?" In: *Acta Neurol Scand* 122.6, pp. 425–9. DOI: 10.1111/j.1600-0404.2010.01333.x.
- Rheenen, W. van et al. (2012). "Hexanucleotide repeat expansions in C9ORF72 in the spectrum of motor neuron diseases". In: *Neurology* 79.9, pp. 878–82. DOI: 10.1212/WNL.0b013e3182661d14.

- Ringholz, G. M., S. H. Appel, M. Bradshaw, N. A. Cooke, D. M. Mosnik, and P. E. Schulz (2005). “Prevalence and patterns of cognitive impairment in sporadic ALS”. In: *Neurology* 65.4, pp. 586–90. DOI: 10.1212/01.wnl.0000172911.39167.b6.
- Rogelj, B. et al. (2012). “Widespread binding of FUS along nascent RNA regulates alternative splicing in the brain”. In: *Sci Rep* 2, p. 603. DOI: 10.1038/srep00603.
- Rosen, D. R. et al. (1993). “Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis”. In: *Nature* 362.6415, pp. 59–62. DOI: 10.1038/362059a0.
- Rossi, D., P. Volanti, L. Brambilla, T. Colletti, R. Spataro, and V. La Bella (2018). “CSF neurofilament proteins as diagnostic and prognostic biomarkers for amyotrophic lateral sclerosis”. In: *J Neurol* 265.3, pp. 510–521. DOI: 10.1007/s00415-017-8730-6.
- Rothstein, J. D., L. J. Martin, and R. W. Kuncel (1992). “Decreased glutamate transport by the brain and spinal cord in amyotrophic lateral sclerosis”. In: *N Engl J Med* 326.22, pp. 1464–8. DOI: 10.1056/NEJM199205283262204.
- Rothstein, J. D., M. Van Kammen, A. I. Levey, L. J. Martin, and R. W. Kuncel (1995). “Selective loss of glial glutamate transporter GLT-1 in amyotrophic lateral sclerosis”. In: *Ann Neurol* 38.1, pp. 73–84. DOI: 10.1002/ana.410380114.
- Rothstein, J. D. et al. (1996). “Knockout of glutamate transporters reveals a major role for astroglial transport in excitotoxicity and clearance of glutamate”. In: *Neuron* 16.3, pp. 675–86.
- Ryberg, H., J. An, S. Darko, J. L. Lustgarten, M. Jaffa, V. Gopalakrishnan, D. Lacomis, M. Cudkowicz, and R. Bowser (2010). “Discovery and verification of amyotrophic lateral sclerosis biomarkers by proteomics”. In: *Muscle Nerve* 42.1, pp. 104–11. DOI: 10.1002/mus.21683.
- Saberi, S., J. E. Stauffer, D. J. Schulte, and J. Ravits (2015). “Neuropathology of Amyotrophic Lateral Sclerosis and Its Variants”. In: *Neurol Clin* 33.4, pp. 855–76. DOI: 10.1016/j.nc1.2015.07.012.

- Sadygov, R. G., D. Cociorva, and J. R. Yates 3rd. (2004). “Large-scale database searching using tandem mass spectra: looking up the answer in the back of the book”. In: *Nat Methods* 1.3, pp. 195–202. DOI: 10.1038/nmeth725.
- Saez, I. and D. Vilchez (2014). “The Mechanistic Links Between Proteasome Activity, Aging and Age-related Diseases”. In: *Curr Genomics* 15.1, pp. 38–51. DOI: 10.2174/138920291501140306113344.
- Sakka, L., G. Coll, and J. Chazal (2011). “Anatomy and physiology of cerebrospinal fluid”. In: *Eur Ann Otorhinolaryngol Head Neck Dis* 128.6, pp. 309–16. DOI: 10.1016/j.anorl.2011.03.002.
- Saman, S. et al. (2012). “Exosome-associated tau is secreted in tauopathy models and is selectively phosphorylated in cerebrospinal fluid in early Alzheimer disease”. In: *J Biol Chem* 287.6, pp. 3842–9. DOI: 10.1074/jbc.M111.277061.
- Sanfilippo, C., A. Longo, F. Lazzara, D. Cambria, G. Distefano, M. Palumbo, A. Cantarella, L. Malaguarnera, and M. Di Rosa (2017). “CHI3L1 and CHI3L2 over-expression in motor cortex and spinal cord of sALS patients”. In: *Mol Cell Neurosci* 85, pp. 162–169. DOI: 10.1016/j.mcn.2017.10.001.
- Sareen, D. et al. (2013). “Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion”. In: *Sci Transl Med* 5.208, 208ra149. DOI: 10.1126/scitranslmed.3007529.
- Sasaki, S. (2011). “Autophagy in spinal cord motor neurons in sporadic amyotrophic lateral sclerosis”. In: *J Neuropathol Exp Neurol* 70.5, pp. 349–59. DOI: 10.1097/NEN.0b013e3182160690.
- Sasaki, S., T. Komori, and M. Iwata (2000). “Excitatory amino acid transporter 1 and 2 immunoreactivity in the spinal cord in amyotrophic lateral sclerosis”. In: *Acta Neuropathol* 100.2, pp. 138–44.
- Sawyers, C. (2004). “Targeted cancer therapy”. In: *Nature* 432.7015, pp. 294–7. DOI: 10.1038/nature03095.

- Saxena, S., E. Cabuy, and P. Caroni (2009). “A role for motoneuron subtype-selective ER stress in disease manifestations of FALS mice”. In: *Nat Neurosci* 12.5, pp. 627–36. DOI: 10.1038/nn.2297.
- Scarafino, A. et al. (2018). “Diagnostic and prognostic power of CSF Tau in amyotrophic lateral sclerosis”. In: *J Neurol* 265.10, pp. 2353–2362. DOI: 10.1007/s00415-018-9008-3.
- Schaeffer-Reiss, C. (2008). “A brief summary of the different types of mass spectrometers used in proteomics”. In: *Methods Mol Biol* 484, pp. 3–16. DOI: 10.1007/978-1-59745-398-1_1.
- Schiffer, D., S. Cordera, P. Cavalla, and A. Migheli (1996). “Reactive astrogliosis of the spinal cord in amyotrophic lateral sclerosis”. In: *J Neurol Sci* 139 Suppl, pp. 27–33.
- Schmidt, A., I. Forne, and A. Imhof (2014). “Bioinformatic analysis of proteomics data”. In: *BMC Syst Biol* 8 Suppl 2, S3. DOI: 10.1186/1752-0509-8-S2-S3.
- Schrooten, M., C. Smetcoren, W. Robberecht, and P. Van Damme (2011). “Benefit of the Awaji diagnostic algorithm for amyotrophic lateral sclerosis: a prospective study”. In: *Ann Neurol* 70.1, pp. 79–83. DOI: 10.1002/ana.22380.
- Sellier, C. et al. (2016). “Loss of C9ORF72 impairs autophagy and synergizes with polyQ Ataxin-2 to induce motor neuron dysfunction and cell death”. In: *EMBO J* 35.12, pp. 1276–97. DOI: 10.15252/embj.201593350.
- Sen, I., A. Nalini, N. B. Joshi, and P. G. Joshi (2005). “Cerebrospinal fluid from amyotrophic lateral sclerosis patients preferentially elevates intracellular calcium and toxicity in motor neurons via AMPA/kainate receptor”. In: *J Neurol Sci* 235.1-2, pp. 45–54. DOI: 10.1016/j.jns.2005.03.049.
- Shah, V., Lassman ME., Chen Y., H. Zhou, and OF. Laterza (2016). “Achieving efficient digestion faster with Flash Digest: potential alternative to multi-step detergent assisted in-solution digestion in quantitative proteomics experiments”. In: *Rapid Commun Mass Spectrom* 31.2, pp. 193–199. DOI: 10.1002/rcm.7778.

- Shahim, P., H. Zetterberg, Y. Tegner, and K. Blennow (2017). “Serum neurofilament light as a biomarker for mild traumatic brain injury in contact sports”. In: *Neurology* 88.19, pp. 1788–1794. DOI: 10.1212/WNL.0000000000003912.
- Shaw, L. M., M. Korecka, C. M. Clark, V. M. Lee, and J. Q. Trojanowski (2007). “Biomarkers of neurodegeneration for diagnosis and monitoring therapeutics”. In: *Nat Rev Drug Discov* 6.4, pp. 295–303. DOI: 10.1038/nrd2176.
- Shaw, P. J., V. Forrest, P. G. Ince, J. P. Richardson, and H. J. Wastell (1995). “CSF and plasma amino acid levels in motor neuron disease: elevation of CSF glutamate in a subset of patients”. In: *Neurodegeneration* 4.2, pp. 209–16.
- Shepherd, S. R., T. Chataway, D. W. Schultz, R. A. Rush, and M. L. Rogers (2014). “The extracellular domain of neurotrophin receptor p75 as a candidate biomarker for amyotrophic lateral sclerosis”. In: *PLoS One* 9.1, e87398. DOI: 10.1371/journal.pone.0087398.
- Shi, M. et al. (2014). “Plasma exosomal alpha-synuclein is likely CNS-derived and increased in Parkinson’s disease”. In: *Acta Neuropathol* 128.5, pp. 639–650. DOI: 10.1007/s00401-014-1314-y.
- Shi, Y. et al. (2018). “Haploinsufficiency leads to neurodegeneration in C9ORF72 ALS/FTD human induced motor neurons”. In: *Nat Med*. DOI: 10.1038/nm.4490.
- Silberring, J. and P. Ciborowski (2010). “Biomarker discovery and clinical proteomics”. In: *Trends Analyt Chem* 29.2, p. 128. DOI: 10.1016/j.trac.2009.11.007.
- Skvortsova, V., M. Shadrina, P. Slominsky, G. Levitsky, E. Kondratieva, A. Zhrebtsova, N. Levitskaya, A. Alekhin, A. Serdyuk, and S. Limborska (2004). “Analysis of heavy neurofilament subunit gene polymorphism in Russian patients with sporadic motor neuron disease (MND)”. In: *Eur J Hum Genet* 12.3, pp. 241–4. DOI: 10.1038/sj.ejhg.5201144.
- Smethurst, P., J. Newcombe, C. Troakes, R. Simone, Y. R. Chen, R. Patani, and K. Sidle (2016). “In vitro prion-like behaviour of TDP-43 in ALS”. In: *Neurobiol Dis* 96, pp. 236–247. DOI: 10.1016/j.nbd.2016.08.007.

- Smith, B. N. et al. (2014). “Exome-wide rare variant analysis identifies TUBA4A mutations associated with familial ALS”. In: *Neuron* 84.2, pp. 324–31. DOI: 10.1016/j.neuron.2014.09.027.
- Smyth, G. K. (2004). “Linear models and empirical bayes methods for assessing differential expression in microarray experiments”. In: *Stat Appl Genet Mol Biol* 3, Article3. DOI: 10.2202/1544-6115.1027.
- Spiller, K. J. et al. (2018). “Microglia-mediated recovery from ALS-relevant motor neuron degeneration in a mouse model of TDP-43 proteinopathy”. In: *Nat Neurosci*. DOI: 10.1038/s41593-018-0083-7.
- Sreedharan, J. et al. (2008). “TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis”. In: *Science* 319.5870, pp. 1668–72. DOI: 10.1126/science.1154584.
- Steenland, K., J. MacNeil, R. Seals, and A. Levey (2010). “Factors affecting survival of patients with neurodegenerative disease”. In: *Neuroepidemiology* 35.1, pp. 28–35. DOI: 10.1159/000306055.
- Steinacker, P. et al. (2018). “Chitotriosidase (CHIT1) is increased in microglia and macrophages in spinal cord of amyotrophic lateral sclerosis and cerebrospinal fluid levels correlate with disease severity and progression”. In: *J Neurol Neurosurg Psychiatry* 89.3, pp. 239–247. DOI: 10.1136/jnnp-2017-317138.
- Steinacker, P. et al. (2016). “Neurofilaments in the diagnosis of motoneuron diseases: a prospective study on 455 patients”. In: *J Neurol Neurosurg Psychiatry* 87.1, pp. 12–20. DOI: 10.1136/jnnp-2015-311387.
- Steinacker, P. et al. (2008). “TDP-43 in cerebrospinal fluid of patients with frontotemporal lobar degeneration and amyotrophic lateral sclerosis”. In: *Arch Neurol* 65.11, pp. 1481–7. DOI: 10.1001/archneur.65.11.1481.
- Storey, John D (2003). “The positive false discovery rate: a Bayesian interpretation and the q-value”. In: *The Annals of Statistics* 31.6, pp. 2013–2035.

- Strong, M. J. and W. Yang (2011). “The frontotemporal syndromes of ALS. Clinicopathological correlates”. In: *J Mol Neurosci* 45.3, pp. 648–55. DOI: 10.1007/s12031-011-9609-0.
- Strong, M. and J. Rosenfeld (2003). “Amyotrophic lateral sclerosis: a review of current concepts”. In: *Amyotroph Lateral Scler Other Motor Neuron Disord* 4.3, pp. 136–43.
- Stuendl, A., M. Kunadt, N. Kruse, C. Bartels, W. Moebius, K. M. Danzer, B. Mollenhauer, and A. Schneider (2016). “Induction of alpha-synuclein aggregate formation by CSF exosomes from patients with Parkinson’s disease and dementia with Lewy bodies”. In: *Brain* 139.Pt 2, pp. 481–94. DOI: 10.1093/brain/awv346.
- Swash, M. (2012). “Why are upper motor neuron signs difficult to elicit in amyotrophic lateral sclerosis?” In: *J Neurol Neurosurg Psychiatry* 83.6, pp. 659–62. DOI: 10.1136/jnnp-2012-302315.
- Swash, M. and M. R. Trimble (2017). “Pseudobulbar affective disorder, emotion and the brain”. In: *J Neurol Neurosurg Psychiatry* 88.10, pp. 809–810. DOI: 10.1136/jnnp-2017-316240.
- Talbot, K., E. Feneberg, J. Scaber, A. G. Thompson, and M. R. Turner (2018). “Amyotrophic lateral sclerosis: the complex path to precision medicine”. In: *J Neurol*. DOI: 10.1007/s00415-018-8983-8.
- Tanaka, M., H. Kikuchi, T. Ishizu, M. Minohara, M. Osoegawa, K. Motomura, T. Tateishi, Y. Ohyagi, and J. Kira (2006). “Intrathecal upregulation of granulocyte colony stimulating factor and its neuroprotective actions on motor neurons in amyotrophic lateral sclerosis”. In: *J Neuropathol Exp Neurol* 65.8, pp. 816–25. DOI: 10.1097/01.jnen.0000232025.84238.e1.
- Tashiro, Y. et al. (2012). “Motor neuron-specific disruption of proteasomes, but not autophagy, replicates amyotrophic lateral sclerosis”. In: *J Biol Chem* 287.51, pp. 42984–94. DOI: 10.1074/jbc.M112.417600.
- Tateishi, T., R. Yamasaki, M. Tanaka, T. Matsushita, H. Kikuchi, N. Isobe, Y. Ohyagi, and J. Kira (2010). “CSF chemokine alterations related to the clinical

- course of amyotrophic lateral sclerosis”. In: *J Neuroimmunol* 222.1-2, pp. 76–81. DOI: 10.1016/j.jneuroim.2010.03.004.
- Taylor, J. P., Jr. Brown R. H., and D. W. Cleveland (2016). “Decoding ALS: from genes to mechanism”. In: *Nature* 539.7628, pp. 197–206. DOI: 10.1038/nature20413.
- Team, R Core (2015). *R: A language and environment for statistical computing*. 3.2.3. Vienna, Austria: R Foundation for Statistical Computing.
- Terada, N., K. Baracskey, M. Kinter, S. Melrose, P. J. Brophy, C. Boucheix, C. Bjartmar, G. Kidd, and B. D. Trapp (2002). “The tetraspanin protein, CD9, is expressed by progenitor cells committed to oligodendrogenesis and is linked to beta1 integrin, CD81, and Tspan-2”. In: *Glia* 40.3, pp. 350–9. DOI: 10.1002/glia.10134.
- Teunissen, C. E. et al. (2011). “Consensus Guidelines for CSF and Blood Biobanking for CNS Biomarker Studies”. In: *Mult Scler Int* 2011, p. 246412. DOI: 10.1155/2011/246412.
- Teyssou, E. et al. (2013). “Mutations in SQSTM1 encoding p62 in amyotrophic lateral sclerosis: genetics and neuropathology”. In: *Acta Neuropathol* 125.4, pp. 511–22. DOI: 10.1007/s00401-013-1090-0.
- Thakore, N. J. and E. P. Piro (2017). “Laughter, crying and sadness in ALS”. In: *J Neurol Neurosurg Psychiatry* 88.10, pp. 825–831. DOI: 10.1136/jnnp-2017-315622.
- Thakur, B. K. et al. (2014). “Double-stranded DNA in exosomes: a novel biomarker in cancer detection”. In: *Cell Res* 24.6, pp. 766–9. DOI: 10.1038/cr.2014.44.
- Therrien, M., G. A. Rouleau, P. A. Dion, and J. A. Parker (2013). “Deletion of C9ORF72 results in motor neuron degeneration and stress sensitivity in *C. elegans*”. In: *PLoS One* 8.12, e83450. DOI: 10.1371/journal.pone.0083450.
- Thery, C., S. Amigorena, G. Raposo, and A. Clayton (2006). “Isolation and characterization of exosomes from cell culture supernatants and biological fluids”.

- In: *Curr Protoc Cell Biol* Chapter 3, Unit 3 22. DOI: 10.1002/0471143030.cb0322s30.
- Thompson, A. G., E. Gray, S. M. Heman-Ackah, I. Mager, K. Talbot, S. E. Andaloussi, M. J. Wood, and M. R. Turner (2016). “Extracellular vesicles in neurodegenerative disease - pathogenesis to biomarkers”. In: *Nat Rev Neurol* 12.6, pp. 346–57. DOI: 10.1038/nrneuro.2016.68.
- Tietje, A., K. N. Maron, Y. Wei, and D. M. Feliciano (2014). “Cerebrospinal fluid extracellular vesicles undergo age dependent declines and contain known and novel non-coding RNAs”. In: *PLoS One* 9.11, e113116. DOI: 10.1371/journal.pone.0113116.
- Tollervey, J. R. et al. (2011). “Characterizing the RNA targets and position-dependent splicing regulation by TDP-43”. In: *Nat Neurosci* 14.4, pp. 452–8. DOI: 10.1038/nn.2778.
- Tomlinson, P. R. et al. (2015). “Identification of distinct circulating exosomes in Parkinson’s disease”. In: *Ann Clin Transl Neurol* 2.4, pp. 353–61. DOI: 10.1002/acn3.175.
- Trajkovic, K., C. Hsu, S. Chiantia, L. Rajendran, D. Wenzel, F. Wieland, P. Schwille, B. Brugger, and M. Simons (2008). “Ceramide triggers budding of exosome vesicles into multivesicular endosomes”. In: *Science* 319.5867, pp. 1244–7. DOI: 10.1126/science.1153124.
- Traynor, B. J., M. B. Codd, B. Corr, C. Forde, E. Frost, and O. Hardiman (1999). “Incidence and prevalence of ALS in Ireland, 1995-1997: a population-based study”. In: *Neurology* 52.3, pp. 504–9.
- Tresse, E., F. A. Salomons, J. Vesa, L. C. Bott, V. Kimonis, T. P. Yao, N. P. Dantuma, and J. P. Taylor (2010). “VCP/p97 is essential for maturation of ubiquitin-containing autophagosomes and this function is impaired by mutations that cause IBMPFD”. In: *Autophagy* 6.2, pp. 217–27.
- Trias, E. et al. (2016). “Post-paralysis tyrosine kinase inhibition with masitinib abrogates neuroinflammation and slows disease progression in inherited amyotrophic

- lateral sclerosis”. In: *J Neuroinflammation* 13.1, p. 177. DOI: 10.1186/s12974-016-0620-9.
- Tsai, G. C., B. Stauch-Slusher, L. Sim, J. C. Hedreen, J. D. Rothstein, R. Kuncel, and J. T. Coyle (1991). “Reductions in acidic amino acids and N-acetylaspartylglutamate in amyotrophic lateral sclerosis CNS”. In: *Brain Res* 556.1, pp. 151–6.
- Tsiatsiani, L. and A. J. Heck (2015). “Proteomics beyond trypsin”. In: *FEBS J* 282.14, pp. 2612–26. DOI: 10.1111/febs.13287.
- Tsuji-Akimoto, S., I. Yabe, M. Niino, S. Kikuchi, and H. Sasaki (2009). “Cystatin C in cerebrospinal fluid as a biomarker of ALS”. In: *Neurosci Lett* 452.1, pp. 52–5.
- Tu, C., P. A. Rudnick, M. Y. Martinez, K. L. Cheek, S. E. Stein, R. J. Slebos, and D. C. Liebler (2010). “Depletion of abundant plasma proteins and limitations of plasma proteomics”. In: *J Proteome Res* 9.10, pp. 4982–91. DOI: 10.1021/pr100646w.
- Turner, M. R. and M. Benatar (2015). “Ensuring continued progress in biomarkers for amyotrophic lateral sclerosis”. In: *Muscle Nerve* 51.1, pp. 14–8. DOI: 10.1002/mus.24470.
- Turner, M. R., A. Cagnin, F. E. Turkheimer, C. C. Miller, C. E. Shaw, D. J. Brooks, P. N. Leigh, and R. B. Banati (2004). “Evidence of widespread cerebral microglial activation in amyotrophic lateral sclerosis: an [11C](R)-PK11195 positron emission tomography study”. In: *Neurobiol Dis* 15.3, pp. 601–9. DOI: 10.1016/j.nbd.2003.12.012.
- Turner, M. R., M. C. Kiernan, P. N. Leigh, and K. Talbot (2009). “Biomarkers in amyotrophic lateral sclerosis”. In: *Lancet Neurol* 8.1, pp. 94–109. DOI: 10.1016/S1474-4422(08)70293-X.
- Turner, M. R. and K. Talbot (2013a). “Mimics and chameleons in motor neurone disease”. In: *Pract Neurol* 13.3, pp. 153–64. DOI: 10.1136/practneurol-2013-000557.

- Turner, M. R. et al. (2013b). “Mechanisms, models and biomarkers in amyotrophic lateral sclerosis”. In: *Amyotroph Lateral Scler Frontotemporal Degener* 14 Suppl 1, pp. 19–32. DOI: 10.3109/21678421.2013.778554.
- Uhlen, M. et al. (2015). “Proteomics. Tissue-based map of the human proteome”. In: *Science* 347.6220, p. 1260419. DOI: 10.1126/science.1260419.
- Vajda, A., R. L. McLaughlin, M. Heverin, O. Thorpe, S. Abrahams, A. Al-Chalabi, and O. Hardiman (2017). “Genetic testing in ALS: A survey of current practices”. In: *Neurology* 88.10, pp. 991–999. DOI: 10.1212/WNL.0000000000003686.
- Valikangas, T., T. Suomi, and L. L. Elo (2018). “A systematic evaluation of normalization methods in quantitative label-free proteomics”. In: *Brief Bioinform* 19.1, pp. 1–11. DOI: 10.1093/bib/bbw095.
- Van Damme, P., W. Robberecht, and L. Van Den Bosch (2017). “Modelling amyotrophic lateral sclerosis: progress and possibilities”. In: *Dis Model Mech* 10.5, pp. 537–549. DOI: 10.1242/dmm.029058.
- Van Deerlin, V. M. et al. (2008). “TARDBP mutations in amyotrophic lateral sclerosis with TDP-43 neuropathology: a genetic and histopathological analysis”. In: *Lancet Neurol* 7.5, pp. 409–16. DOI: 10.1016/S1474-4422(08)70071-1.
- Van Den Bosch, L., W. Vandenberghe, H. Klaassen, E. Van Houtte, and W. Robberecht (2000). “Ca(2+)-permeable AMPA receptors and selective vulnerability of motor neurons”. In: *J Neurol Sci* 180.1-2, pp. 29–34.
- Van Langenhove, T. et al. (2010). “Genetic contribution of FUS to frontotemporal lobar degeneration”. In: *Neurology* 74.5, pp. 366–71. DOI: 10.1212/WNL.0b013e3181ccc732.
- Varghese, A. M., A. Sharma, P. Mishra, K. Vijayalakshmi, H. C. Harsha, T. N. Sathyaprabha, S. M. Bharath, A. Nalini, P. A. Alladi, and T. R. Raju (2013). “Chitotriosidase - a putative biomarker for sporadic amyotrophic lateral sclerosis”. In: *Clin Proteomics* 10.1, p. 19. DOI: 10.1186/1559-0275-10-19.

- Vella, L. J., R. A. Sharples, V. A. Lawson, C. L. Masters, R. Cappai, and A. F. Hill (2007). “Packaging of prions into exosomes is associated with a novel pathway of PrP processing”. In: *J Pathol* 211.5, pp. 582–90. DOI: 10.1002/path.2145.
- Vilchez, D., I. Saez, and A. Dillin (2014). “The role of protein clearance mechanisms in organismal ageing and age-related diseases”. In: *Nat Commun* 5, p. 5659. DOI: 10.1038/ncomms6659.
- Visser, A. E. et al. (2018). “Multicentre, cross-cultural, population-based, case-control study of physical activity as risk factor for amyotrophic lateral sclerosis”. In: *J Neurol Neurosurg Psychiatry*. DOI: 10.1136/jnnp-2017-317724.
- Vu, L. T. and R. Bowser (2017). “Fluid-Based Biomarkers for Amyotrophic Lateral Sclerosis”. In: *Neurotherapeutics* 14.1, pp. 119–134. DOI: 10.1007/s13311-016-0503-x.
- Vucic, S., G. A. Nicholson, and M. C. Kiernan (2008). “Cortical hyperexcitability may precede the onset of familial amyotrophic lateral sclerosis”. In: *Brain* 131.Pt 6, pp. 1540–50. DOI: 10.1093/brain/awn071.
- Wadia, J. S., M. Schaller, R. A. Williamson, and S. F. Dowdy (2008). “Pathologic prion protein infects cells by lipid-raft dependent macropinocytosis”. In: *PLoS One* 3.10, e3314. DOI: 10.1371/journal.pone.0003314.
- Walker, A. K., K. J. Spiller, G. Ge, A. Zheng, Y. Xu, M. Zhou, K. Tripathy, L. K. Kwong, J. Q. Trojanowski, and V. M. Lee (2015). “Functional recovery in new mouse models of ALS/FTLD after clearance of pathological cytoplasmic TDP-43”. In: *Acta Neuropathol* 130.5, pp. 643–60. DOI: 10.1007/s00401-015-1460-x.
- Wang, H., M. Wu, C. Zhan, E. Ma, M. Yang, X. Yang, and Y. Li (2012). “Neurofilament proteins in axonal regeneration and neurodegenerative diseases”. In: *Neural Regen Res* 7.8, pp. 620–6. DOI: 10.3969/j.issn.1673-5374.2012.08.010.
- Wang, M. D., J. Little, J. Gomes, N. R. Cashman, and D. Krewski (2017). “Identification of risk factors associated with onset and progression of amyotrophic lateral sclerosis using systematic review and meta-analysis”. In: *Neurotoxicology* 61, pp. 101–130. DOI: 10.1016/j.neuro.2016.06.015.

- Webb-Robertson, B. J. et al. (2015). “Review, evaluation, and discussion of the challenges of missing value imputation for mass spectrometry-based label-free global proteomics”. In: *J Proteome Res* 14.5, pp. 1993–2001. DOI: 10.1021/pr501138h.
- Welton, J. L., S. Loveless, T. Stone, C. von Ruhland, N. P. Robertson, and A. Clayton (2017). “Cerebrospinal fluid extracellular vesicle enrichment for protein biomarker discovery in neurological disease; multiple sclerosis”. In: *J Extracell Vesicles* 6.1, p. 1369805. DOI: 10.1080/20013078.2017.1369805.
- Wessel, D. and U. I. Flugge (1984). “A method for the quantitative recovery of protein in dilute solution in the presence of detergents and lipids”. In: *Anal Biochem* 138.1, pp. 141–3.
- Westeneng, H. J. et al. (2018). “Prognosis for patients with amyotrophic lateral sclerosis: development and validation of a personalised prediction model”. In: *Lancet Neurol* 17.5, pp. 423–433. DOI: 10.1016/S1474-4422(18)30089-9.
- Weydt, P. et al. (2016). “Neurofilament levels as biomarkers in asymptomatic and symptomatic familial amyotrophic lateral sclerosis”. In: *Ann Neurol* 79.1, pp. 152–8. DOI: 10.1002/ana.24552.
- Wilke, C., C. Deuschle, T. W. Rattay, W. Maetzler, and M. Synofzik (2015). “Total tau is increased, but phosphorylated tau not decreased, in cerebrospinal fluid in amyotrophic lateral sclerosis”. In: *Neurobiol Aging* 36.2, pp. 1072–4. DOI: 10.1016/j.neurobiolaging.2014.10.019.
- Wilson, M. E., I. Boumaza, D. Lacomis, and R. Bowser (2010). “Cystatin C: a candidate biomarker for amyotrophic lateral sclerosis”. In: *PLoS One* 5.12, e15133. DOI: 10.1371/journal.pone.0015133.
- Winer, L. et al. (2013). “SOD1 in cerebral spinal fluid as a pharmacodynamic marker for antisense oligonucleotide therapy”. In: *JAMA Neurol* 70.2, pp. 201–7. DOI: 10.1001/jamaneurol.2013.593.
- Wuolikainen, A., T. Moritz, S. L. Marklund, H. Antti, and P. M. Andersen (2011). “Disease-related changes in the cerebrospinal fluid metabolome in amyotrophic

- lateral sclerosis detected by GC/TOFMS”. In: *PLoS One* 6.4, e17947. DOI: 10.1371/journal.pone.0017947.
- Xiao, L., D. Ohayon, I. A. McKenzie, A. Sinclair-Wilson, J. L. Wright, A. D. Fudge, B. Emery, H. Li, and W. D. Richardson (2016). “Rapid production of new oligodendrocytes is required in the earliest stages of motor-skill learning”. In: *Nat Neurosci* 19.9, pp. 1210–1217. DOI: 10.1038/nn.4351.
- Yadav, S., L. G. Carrascosa, A. A. Sina, M. J. Shiddiky, M. M. Hill, and M. Trau (2016). “Electrochemical detection of protein glycosylation using lectin and protein-gold affinity interactions”. In: *Analyst* 141.8, pp. 2356–61. DOI: 10.1039/c6an00528d.
- Yamashita, S. and Y. Ando (2015). “Genotype-phenotype relationship in hereditary amyotrophic lateral sclerosis”. In: *Transl Neurodegener* 4, p. 13. DOI: 10.1186/s40035-015-0036-y.
- Yang, M. L., J. Shin, C. A. Kearns, M. M. Langworthy, H. Snell, M. B. Walker, and B. Appel (2015). “CNS myelination requires cytoplasmic dynein function”. In: *Dev Dyn* 244.2, pp. 134–45. DOI: 10.1002/dvdy.24238.
- Youden, W. J. (1950). “Index for rating diagnostic tests”. In: *Cancer* 3.1, pp. 32–5.
- Zetterberg, H., J. Jacobsson, L. Rosengren, K. Blennow, and P. M. Andersen (2007). “Cerebrospinal fluid neurofilament light levels in amyotrophic lateral sclerosis: impact of SOD1 genotype”. In: *Eur J Neurol* 14.12, pp. 1329–33. DOI: 10.1111/j.1468-1331.2007.01972.x.
- Zetterstrom, P., P. M. Andersen, T. Brannstrom, and S. L. Marklund (2011). “Misfolded superoxide dismutase-1 in CSF from amyotrophic lateral sclerosis patients”. In: *J Neurochem* 117.1, pp. 91–9. DOI: 10.1111/j.1471-4159.2011.07177.x.
- Zhang, Y., B. R. Fonslow, B. Shan, M. C. Baek, and 3rd Yates J. R. (2013). “Protein analysis by shotgun/bottom-up proteomics”. In: *Chem Rev* 113.4, pp. 2343–94. DOI: 10.1021/cr3003533.

- Zheng, W., S. A. Johnston, and L. Joshua-Tor (1998). “The unusual active site of Gal6/bleomycin hydrolase can act as a carboxypeptidase, aminopeptidase, and peptide ligase”. In: *Cell* 93.1, pp. 103–9.
- Zhu, H. et al. (2001). “Global analysis of protein activities using proteome chips”. In: *Science* 293.5537, pp. 2101–5. DOI: 10.1126/science.1062191.
- Zou, Z. Y., Z. R. Zhou, C. H. Che, C. Y. Liu, R. L. He, and H. P. Huang (2017). “Genetic epidemiology of amyotrophic lateral sclerosis: a systematic review and meta-analysis”. In: *J Neurol Neurosurg Psychiatry* 88.7, pp. 540–549. DOI: 10.1136/jnnp-2016-315018.
- Zurcher, N. R. et al. (2015). “Increased in vivo glial activation in patients with amyotrophic lateral sclerosis: assessed with [(11)C]-PBR28”. In: *Neuroimage Clin* 7, pp. 409–14. DOI: 10.1016/j.nicl.2015.01.009.