

17 **A. Flat Files**

18

19 **Complete the Inventory below for all additional textual information**  
 20 **and any additional Supplementary Figures, which should be supplied in**  
 21 **one combined PDF file.**

22

23 • **Row 1:** A combined, flat PDF containing any Supplementary Text, Discussion,  
 24 Notes, Additional Supplementary Figures, Supplementary Protocols, simple tables,  
 25 and all associated legends. Only one such file is permitted.

26

27 • **Row 2:** Nature Research’s Reporting Summary; if previously requested by the  
 28 editor, please provide an updated Summary, fully completed, without any mark-  
 29 ups or comments. **(Reporting Summaries are not required for all**  
 30 **manuscripts.)**

31

Item	Present?	Filename This should be the name the file is saved as when it is uploaded to our system, and should include the file extension. The extension must be .pdf	A brief, numerical description of file contents. i.e.: <i>Supplementary Figures 1-4, Supplementary Discussion, and Supplementary Tables 1-4.</i>
Supplementary Information	No		
Reporting Summary	Yes	nr-reporting-summary_NME D-L102212B.pdf	

32

33 **Editor summary:** Increased perivascular fibroblast activity and vascular remodeling occurs  
 34 early in sporadic ALS pathogenesis and can predict patient survival time

35 **Reviewer recognition statement:** Nature Medicine thanks Robert Baloh and the other,  
 36 anonymous, reviewers for their contribution to the peer review of this work.

37

38

39 **Title:** *Altered perivascular fibroblast activity precedes ALS disease onset.*

40

41 **Authors:** Anna Månberg<sup>1#</sup>, Nathan Skene<sup>2,3,4#</sup>, Folkert Sanders<sup>5</sup>, Marta Trusohamn<sup>2</sup>, Julia Remnestål<sup>1</sup>,  
 42 Anna Szczepińska<sup>5</sup>, Inci Sevval Aksoylu<sup>5</sup>, Peter Lönnerberg<sup>2</sup>, Lwaki Ebarasi<sup>6</sup>, Stefan Wouters<sup>5</sup>, Manuela  
 43 Lehmann<sup>7</sup>, Jennie Olofsson<sup>1</sup>, Inti Von Gohren Antequera<sup>5</sup>, Aylin Domaniku<sup>5</sup>, Maxim De Schaepdryver<sup>8</sup>,

44 Joke De Vocht<sup>9</sup>, Koen Poesen<sup>8,10</sup>, Mathias Uhlén<sup>11,12</sup>, Jasper Anink<sup>13</sup>, Caroline Mijnsbergen<sup>13</sup>,  
45 Hermieneke Vergunst-Bosch<sup>14</sup>, Annemarie Hübers<sup>15,16</sup>, Ulf Kläppe<sup>17</sup>, Elena Rodriguez-Vieitez<sup>18</sup>  
46 Jonathan D. Gilthorpe<sup>7</sup>, Eva Hedlund<sup>12</sup>, Robert A. Harris<sup>5</sup>, Eleonora Aronica<sup>13</sup>, Philip Van Damme<sup>9</sup>,  
47 Albert Ludolph<sup>15,19</sup>, Jan Veldink<sup>14</sup>, Caroline Ingre<sup>17</sup>, Peter Nilsson<sup>1</sup>, Sebastian A. Lewandowski<sup>1,5\*</sup>

48

49 **Affiliations and research teams:**

50 <sup>1</sup> Div. of Affinity Proteomics, Dept. of Protein Science, KTH Royal Institute of Technology, SciLifeLab,  
51 Stockholm, Sweden

52 <sup>2</sup> Dept. of Medical Biochemistry and Biophysics, Karolinska Institute, Stockholm, Sweden

53 <sup>3</sup> Division of Neuroscience, Dept. of Brain Sciences, Imperial College London, London, United Kingdom

54 <sup>4</sup> United Kingdom Dementia Research Institute, London, United Kingdom

55 <sup>5</sup> Karolinska Institute, Dept. of Clinical Neuroscience, Centre for Molecular Medicine, Karolinska  
56 Hospital, Stockholm, Sweden

57 <sup>6</sup> Dept. of Laboratory Medicine, Karolinska Institute, Stockholm, Sweden

58 <sup>7</sup> Umeå University, Dept. of Integrative Medical Biology, Umeå, Sweden

59 <sup>8</sup> Laboratory for Neurobiomarker Research, Dept. of Neurology, Leuven Brain Institute, KU Leuven  
60 (University of Leuven), Leuven, Belgium

61 <sup>9</sup> KU Leuven, Neurology Dept. and Center for Brain & Disease Research, VIB, Leuven, Belgium

62 <sup>10</sup> Laboratory Medicine, UZ Leuven (University Hospital Leuven), Leuven, Belgium

63 <sup>11</sup> Div. Systems Biology, Dept. Protein Science, KTH Royal Institute of Technology, SciLifeLab,  
64 Stockholm, Sweden

65 <sup>12</sup> Dept. Neuroscience, Karolinska Institute, Stockholm, Sweden

66 <sup>13</sup> Amsterdam UMC, University of Amsterdam, Dept. of (Neuro)Pathology, Amsterdam Neuroscience,  
67 Amsterdam, The Netherlands

68 <sup>14</sup> UMC Utrecht Brain Center, University Medical Center Utrecht, Dept. of Neurology, Utrecht  
69 University, Utrecht, The Netherlands

70 <sup>15</sup> University of Ulm, Neurology Clinic, Ulm, Germany

71 <sup>16</sup> Current affiliation: Neurology Division, Geneva University Hospital.

72 <sup>17</sup> Dept. of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden. Dept of Neurology at  
73 Karolinska University Hospital.

74 <sup>18</sup> Division of Clinical Geriatrics, Department of Neurobiology, Care Sciences and Society, Karolinska  
75 Institutet, Stockholm, Sweden

76 <sup>19</sup> Deutsches Zentrum für Neurodegenerative Erkrankungen (DZNE), Ulm

77

78 # These authors contributed equally.

79 \* Corresponding author: sebastian.lewandowski@ki.se

80 **Abstract:**

81 Apart from the well-defined factors in neuronal cells<sup>1</sup>, only few reports consider that variability of  
82 sporadic ALS progression can depend on the less-defined contributions from glia<sup>2,3</sup> and blood vessels<sup>4</sup>.  
83 In this study we use an expression weighted cell-type enrichment method to infer cell activity in spinal  
84 cord samples from sporadic ALS patients and mouse models of this disease. Here we report that  
85 sporadic ALS patients present cell activity patterns consistent with two mouse models in which  
86 enrichments of vascular cell genes preceded the microglial response. Notably, during the  
87 presymptomatic stage, perivascular fibroblast cells showed the strongest gene enrichments and their  
88 marker proteins SPP1 and COL6A1 accumulated in enlarged perivascular spaces in sporadic ALS  
89 patients. Moreover, in plasma of 574 ALS patients from four independent cohorts, increased levels of  
90 SPP1 at disease diagnosis repeatedly predicted shorter survival with stronger effect than the  
91 established risk factors of bulbar onset or neurofilament levels in cerebrospinal fluid. We propose that  
92 the activity of the recently-discovered perivascular fibroblast can predict ALS patient survival and  
93 provide a novel conceptual framework to re-evaluate definitions of ALS etiology.

94 **Introduction**

95 Although the defining clinical features of amyotrophic lateral sclerosis (ALS) focus on the  
96 common degeneration of upper and lower motor neurons, substantial variability is reported for  
97 patient's age at disease onset, contribution of non-motor systems and survival duration<sup>5</sup>. This  
98 variability considerably complicates accurate prognosis of life expectancy and interpretations of  
99 clinical trial outcomes. Lack of reliable early diagnostics means that treatments are given only in  
100 advanced stages; and further highlights that studies aiming to understand early sporadic disease  
101 mechanisms are intrinsically challenging and incomplete<sup>6</sup>.

102 From the first description of familial ALS forms<sup>7</sup> which affect 5-10% of patients, there have  
103 been remarkable advances in discovery of novel gene variants. These variants provided insight into  
104 dysfunction of neuron-centric pathological mechanisms including protein misfolding<sup>7</sup>, RNA maturation<sup>8</sup>  
105 and axonal transport<sup>9</sup>. However, increasing evidence indicates that the dynamics of  
106 neurodegeneration in both familial and sporadic ALS is also influenced by other cell type functions that  
107 include glia response<sup>2,3</sup>, oligodendrocyte metabolism<sup>10</sup> and integrity of blood vessels<sup>4,11</sup>. These  
108 observations challenged the neuron-centric theory and initiated redefinition of our understanding of  
109 disease variability. Recent efforts to comprehend cellular complexity of the nervous system using  
110 single cell RNA sequencing provided detailed maps of novel cell types<sup>12-14</sup> but also add another level of  
111 challenge to interpret individual cell-type input during multiple stages of disease. Unless we improve  
112 our understanding of contributions from non-neuronal cells and mechanisms preceding disease onset,  
113 the clinical variability of the progression dynamics will continue to confound design and evaluation of  
114 ALS treatments. We therefore aimed to decode the temporal activity for gene expression within ten  
115 major central nervous system cell types in post-mortem transcriptomes of patients with sporadic ALS  
116 and in a presymptomatic timeline of the transgenic *SOD1*<sup>G93A</sup> and *TARDBP*<sup>Q331K/Q331K</sup> mouse models.

117

118 **Cell activity timeline in ALS**

119 We used single-cell central nervous system transcriptomes<sup>12,13</sup> and an expression weighted cell-type  
120 enrichment (EWCE)<sup>15</sup> method to infer the cell activity in bulk tissue. This form of inferred activity is  
121 represented as a z-score which measures the degree of transcript increase and specificity within a  
122 given cell category derived from single-cell sequencing. In our analysis we included transcriptomes  
123 from spinal cords of sporadic ALS patients (n=12) and two transgenic mouse models: *SOD1*<sup>G93A</sup> (n=3)  
124 and *TARDBP*<sup>Q331K/Q331K</sup> (n≥8) (Fig. 1 a, b Extended Data Fig. 1, 2). The results from sporadic ALS patient  
125 tissues revealed increased enrichments of microglia and astrocyte-specific genes together with a  
126 marked reduction of excitatory neuron and interneuron transcripts likely reflecting neurodegeneration  
127 (Fig. 1 b). Notably we observed increased activity of vascular cell-specific genes in sporadic ALS  
128 patients with the highest enrichments for perivascular fibroblast cells (Fig. 1 b). To complement the  
129 endpoint patient data and to explore presymptomatic patterns of cell activity we analyzed gene  
130 enrichments for the same ten cell types in *SOD1*<sup>G93A</sup> and *TARDBP*<sup>Q331K/Q331K</sup> mice. We analyzed *SOD1*<sup>G93A</sup>

131 mice at ages of 4, 6 (asymptomatic); 8 (neuromuscular junction decoupling<sup>16</sup>); 10 (pre-onset); 14  
132 (onset of neuroinflammation<sup>17</sup>); 16 (peak body weight - clinical onset with loss of motor neuron  
133 count<sup>18</sup>) and 18 weeks (symptomatic stage). *TARDBP*<sup>Q331K/Q331K</sup> mice were analyzed at 5 month (no MN  
134 loss or NMJ denervation with partial behavior phenotype<sup>19</sup>) and 20 month (with behavioral phenotype)  
135 timepoints. We observed a remarkable consistency in cell type activity between end-stage sporadic  
136 ALS patients and symptomatic *SOD1*<sup>G93A</sup> or *TARDBP*<sup>Q331K/Q331K</sup> mice with similar increased expression of  
137 microglia and astrocyte associated genes as well as a reduction of neuron and interneuron transcripts  
138 (Fig. 1b, Extended Data Fig. 1). Our observations in *SOD1*<sup>G93A</sup> mice were consistent with previously  
139 published FACS sorting<sup>20</sup> and RNA pulldown<sup>21</sup> approaches, which also aimed to identify cell type-  
140 specific transcripts. For instance, we identified *Trem2*, *Ctss* with *Cd86* as well as *Sprr1a* with *Klk6* as  
141 respective microglia and oligodendrocyte-specific regulated genes, similarly to the previous reports  
142 from in *SOD1*<sup>G93A</sup><sup>20</sup> and *SOD1*<sup>G37R</sup><sup>21</sup> models (Extended Data Fig. 1). In comparison with those reports,  
143 our dataset not only includes novel cell types, but also resolves with a more refined timescale to  
144 demonstrate that microglia and astrocytes become induced at pre-onset week 14, while neuron and  
145 interneuron genes already show decreased activity in presymptomatic 8 week-old mice.

146 Remarkably, in the *SOD1*<sup>G93A</sup> model the response of vascular cells including pericyte- (4-6  
147 weeks), smooth muscle- (8 weeks) and endothelial cell-specific genes (10 weeks) largely occurred  
148 before the microglial response (week 14-18) and showed strongest enrichment for perivascular  
149 fibroblast-specific genes (4, 8, 14, 16 and 18 weeks) (Fig. 1b and Extended Data Fig. 2a). A similar  
150 increase of perivascular fibroblast-specific genes was also evident and preceded microglial  
151 transcriptional activity in 5 month-old *TARDBP*<sup>Q331K/Q331K</sup> mice (Fig. 1 b). The timing and degree of  
152 perivascular fibroblast-specific gene expression could imply that the extent of their activity is similar to  
153 microglia, but they become induced at much earlier stages of disease. We therefore focused our  
154 analysis on the perivascular fibroblast-specific expression patterns and defined several activated genes  
155 in human sporadic ALS patients as well as in *SOD1*<sup>G93A</sup> and *TARDBP*<sup>Q331K/Q331K</sup> mice (Fig. 1c, d, e). The  
156 timing of expression for perivascular fibroblast-specific genes in *SOD1*<sup>G93A</sup> mice (Fig. 1e) implies that  
157 one group becomes active during the presymptomatic stage of 4-8 weeks, while another group  
158 become active together with the microglial response initiated at week 14. We found similar patterns of  
159 early and late responses of perivascular fibroblast-specific genes in *TARDBP*<sup>Q331K/Q331K</sup> mice (Fig. 1 e).

160 Perivascular fibroblasts were recently identified as a unique nervous system cell type by other  
161 groups using single-cell RNA sequencing efforts<sup>13,22</sup> (Fig. 1f) and are also referred to as vascular  
162 leptomeningeal cells<sup>13,14</sup>. These cells were previously reported to locate within the perivascular  
163 Virchow–Robin space between the mural cells and astrocyte end-feet<sup>22</sup> (Fig. 1g). Our transmission  
164 electron microscopy observations further specify the location of perivascular fibroblasts between the  
165 astrocyte and mural cell basement membranes (Fig. 1h). Perivascular fibroblast-specific genes (e.g.  
166 *Col6a1*, *Col1a1* and *Mmp2*)<sup>13,22</sup> imply that their main functions include composition and remodeling of  
167 the basement membrane extracellular matrix<sup>22</sup>. However, as current tools are inadequate to perform

168 recombinant cell-specific interventions, the precise role of these cells in the nervous system remains  
169 highly debated. To enable the exploration of our dataset we have built a web resource with cell type-  
170 specific enrichment rankings, transcript specificity and expression dynamics in sporadic ALS patients,  
171 *SOD1<sup>G93A</sup>* and *TARDBP<sup>Q331K/Q331K</sup>* mice at <http://alscellatlas.org> ).

172

### 173 **Vascular remodeling in ALS**

174 Since perivascular fibroblast cells in our data were the first to induce gene expression before  
175 ALS onset and their function is largely unknown, we focused on addressing their potential  
176 contributions to disease processes. Our computational predictions were based on RNA expression  
177 specificity in a healthy state and we needed to verify if perivascular fibroblast protein expression  
178 indeed localized within blood vessels in ALS tissues. We performed a histological screen with  
179 antibodies targeting 15 perivascular fibroblast-specific proteins. The analysis confirmed vascular  
180 location for several proteins, in particular COL6A1 and SPP1 (osteopontin) (Fig. 2a, b and Extended  
181 Data Fig. 3). Stainings of both proteins were increased in spinal cords of sporadic ALS patients as  
182 compared with age-matched non-neurological controls (Fig. 2c, Extended Data Fig. 3a). In addition,  
183 vessel-specific increase of Spp1 and Col6a1 protein expression in *SOD1<sup>G93A</sup>* mice spinal cords (Fig. 2d)  
184 was consistent with those observed in sporadic ALS patients. Accumulation of both proteins already  
185 occurred at an early timepoint of 8 weeks (Fig. 2e), which supports the inferred notion from the RNA  
186 transcriptomics data that perivascular fibroblasts become induced at presymptomatic stage of ALS.

187 Notably, staining for COL6A1, a component of the vascular basement membrane, did not  
188 outline endothelial or mural cells, but indicated enlarged perivascular spaces in ALS patients (Fig. 2b  
189 and Extended Data Fig. 3a) suggesting that astrocyte and mural basement membranes become  
190 separated during ALS progression. We confirmed that separation of mural basement membrane  
191 (marked with COL4A1) and astrocyte basement membrane (marked with LAMA1) does indeed occur in  
192 spinal cord vessels in sporadic ALS patients (Extended Data Fig. 3b). A similar separation of Col4a1 and  
193 Lama1-delineated basement membranes already occurred in presymptomatic (8 week-old) *SOD1<sup>G93A</sup>*  
194 mice (Fig. 2f) at a timepoint consistent with increased transcriptional activity and protein accumulation  
195 of perivascular fibroblast-specific markers. We resolved the ultrastructure of basement membrane  
196 separation using transmission electron microscopy and observed that perivascular fibroblasts were  
197 present within enlarged perivascular spaces in 14 week-old *SOD1<sup>G93A</sup>* mice (Fig. 2g). We found that  
198 these perivascular spaces accumulate perivascular fibroblast-specific COL6A1 and SPP1 proteins in  
199 spinal cord vessels of sporadic ALS patients (Fig. 2h, i and Extended Data Fig. 4a). The enlargement of  
200 perivascular spaces was substantial in both end-stage ALS patients and in 14 week-old *SOD1<sup>G93A</sup>* mice  
201 (Fig. 2j). This form of basement membrane remodeling is likely regulated by multiple genes, as  
202 indicated by enriched gene ontologies which occurred early and were exacerbated in symptomatic  
203 stages of ALS (Extended Data Fig. 5). Taken together, these findings indicate that perivascular  
204 fibroblasts become active during presymptomatic ALS stages and can participate in remodeling of

205 cerebral blood vessels by increased expression of the COL6A1 and SPP1 proteins within enlarged  
206 perivascular spaces (Fig. 2k).

207

### 208 **Prognostic value of perivascular fibroblast proteins in plasma**

209 Since perivascular spaces are responsible for clearance from cerebrospinal fluid to lymphatic<sup>23</sup> and  
210 blood<sup>24</sup> circulations, we anticipated that an increase of COL6A1 and SPP1 in ALS patient plasma could  
211 have potential informative value for clinical disease dynamics. We performed targeted suspension  
212 bead array profiling using 32 antibodies against 15 perivascular fibroblast proteins in blood plasma.  
213 Initial screening revealed a that higher quartile values of COL6A1 and SPP1 proteins indicated shorter  
214 survival of ALS patients in a pilot cohort from the Netherlands (189 ALS patients and 199 controls)  
215 (Extended Data Fig. 7a). We have combined the Netherlands cohort with groups from Germany and  
216 Belgium (a total of 452 ALS patients and 395 controls) and analyzed it as a discovery cohort with  
217 internal replication between countries. We have also recruited samples from Sweden (122 ALS  
218 patients and 109 controls) as a replication cohort with independent labeling of beads with antibodies.  
219 Cohort summary statistics are presented in Extended Data Fig. 6.

220 SPP1 levels above empirically optimized thresholds indicated shorter survival in the discovery  
221 and replication cohorts (Fig. 3a). These associations of high SPP1 and shorter survival were reproduced  
222 within the individual cohorts from Netherlands and Germany but not in the smaller cohort from  
223 Belgium with less complete survival data (Extended Data Fig. 8). We also performed a validation using  
224 a second SPP1 antibody targeting an independent protein domain (Extended Data Fig. 7b). This  
225 antibody showed similar association between high SPP1 values and short patient survival in each  
226 country cohort (Extended Data Fig. 7, 8b, 9a). Plasma COL6A1 level increase also associated with  
227 shorter survival in the grouped discovery cohort and within the larger groups from Netherlands  
228 (n=189) and Germany (n=187), but not in the smaller cohorts with less complete survival data from  
229 Belgium (n=102) and Sweden (n=122) (Extended Data Fig. 8c, 9a).

230 Besides the comparison of patient groups defined by protein level threshold, we also  
231 evaluated the prognostic risk for continuous increase of plasma SPP1 and COL6A1 in univariate and  
232 multivariate Cox models. We compared their hazard ratios to the effects of bulbar onset and, when  
233 available, neurofilament light (NFL) levels. Using a univariate model, the hazard ratio for continuous  
234 plasma increase of SPP1 in the grouped discovery cohort was higher (HR=1.82 p=8.96e-05) than the  
235 risk mediated by the bulbar onset (HR=1.44 p=2.1e-05) (Fig. 3c). This effect of univariate risk increase  
236 was consistent in the Swedish replication cohort in which SPP1 increase had higher hazard ratio  
237 (HR=3.27 p=0.0008) than bulbar onset (HR=1.66 p=0.0378) or NFL increase in CSF (HR=1.61 p=0.002)  
238 (Fig. 3c). We have further confirmed these observations using the second antibody for SPP1 and  
239 univariate Cox models in discovery and replication cohorts (Extended Data Fig. 9b). In addition to the  
240 univariate model, we cross-examined the effects of continuous protein increase using a multivariate  
241 Cox model in which SPP1 persisted to give stronger survival hazard ratios (HR=1.82 p=8.96e-05) than

242 the bulbar onset in the discovery cohort (Fig. 3c). Strikingly, in the replication cohort where matched  
243 CSF samples were available, the multivariate effect of SPP1 increase in plasma had stronger indication  
244 for short survival (HR=3.56 p=0.009) than did the NFL increase in CSF as measured using a clinical grade  
245 ELISA kit (HR=1.5 p=0.02) (Fig. 3c). The increase of SPP1 in CSF also indicated shorter survival in the  
246 Swedish cohort, albeit at much higher threshold and to a lesser extent than did the plasma values  
247 (Extended data Fig. 9c). The correlations between SPP1 and COL6A1 were more apparent in CSF than  
248 in plasma (Extended Data Fig. 10a), which could support the proposed CNS origin of these proteins as  
249 indicated by our transcriptomics and histological data. However, the relative and longitudinal increase  
250 of SPP1 was more evident in ALS patient plasma rather than in CSF (Extended data Fig. 10 b, c) which  
251 was concordant with a higher predictive value in plasma.

252

## 253 **Discussion**

254 Our findings indicate the interdependence of vascular and neuronal systems and illustrate the  
255 novel, early contributions from perivascular fibroblast cells which were not previously considered in  
256 interpretations of ALS etiology. We propose that inclusion of vascular cell-specific inputs will likely  
257 improve prognostic accuracy at disease diagnosis and can help to re-evaluate the current view on  
258 mechanisms that facilitate neurodegeneration. Although the current neuron-centric view is valid from  
259 the motor phenotype standpoint, it may not be sufficient to explain the reported variety of sporadic  
260 ALS phenotypes<sup>5,25</sup> and disease dynamics<sup>26</sup>. The existing assessments of disease dynamics, which are  
261 often based on neuronal axon-derived inputs (e.g. neurofilaments), general epidemiological  
262 descriptions (e.g. age at onset, diagnostic delay) or physiological factors (e.g. bulbar onset, functional  
263 rating scales), provide limited perspective for at least two reasons.

264 Firstly, current prognostic and causative assumptions are based on indicators of outcome (e.g.  
265 neurofilament proteins and disease progression scales) which likely reflect the damage already  
266 occurred, but not necessarily the early mechanisms and cell types that precede symptom onset. These  
267 presymptomatic mechanisms may not be directly reflected by motor phenotypes, as indicated by our  
268 mouse model observations, and are intrinsically difficult to observe in sporadic ALS patients. In  
269 addition, a recent attempt to stage ALS has defined its neuronal onset in the motor cortex but, in  
270 contrast to Alzheimer's and Parkinson's disease, showed no preclinical stages for neuronal cells<sup>27</sup>. Our  
271 report therefore focuses on cellular activity in presymptomatic ALS mice and demonstrates that  
272 induction of perivascular fibroblast transcripts, and deposition of the respective proteins largely  
273 precede neuroinflammation<sup>17</sup> and neuronal loss<sup>28</sup>.

274 Secondly, considering the interdependency of cell types in the nervous system, the net  
275 outcome for ALS neurodegeneration dynamics is not only associated with, but also relies on support  
276 mechanisms from multiple cell types including blood vessels. However, current molecular factors that  
277 indicate survival variability in the clinics are mostly based on neuron-derived proteins and only  
278 recently on concurrent microglia-derived<sup>29</sup> proteins which reflect late symptomatic stages of disease.



279 Input from these cell types likely represents only partial accuracy of early prognosis and limits the  
280 design of therapeutic or preventive clinical trials. Herein we showed that the SPP1 protein alone had a  
281 stronger indication on survival prediction than did standard physiological or neuron-based inputs when  
282 measured at first clinical visit. Moreover, the increase of SPP1 remained the strongest indicator for  
283 poor survival when compared against bulbar onset and neurofilament levels in CSF using multivariate  
284 models.

285 We thus propose that the vascular response is a major independent indicator for ALS patient  
286 survival. Even though vascular dysfunction has previously been proposed as an epidemiological risk  
287 factor in humans<sup>30,31</sup> and is observed in animal models<sup>4,11</sup> our results instead provide a plausible  
288 molecular explanation for these associations, a potential target for therapy as well as practical  
289 prognostic tools for the wide population of sporadic ALS patients. Although speculative at this point,  
290 vascular remodeling may be important to ALS neurodegeneration because of its reported associated  
291 effects on the blood-brain barrier function<sup>4</sup>, reduced blood flow<sup>32,33</sup>, and/or decreased cerebral  
292 glucose uptake<sup>34</sup>. The prognostic value of SPP1 levels likely represents a general vascular injury  
293 response and has been reported beyond ALS neurodegeneration in several publications showing an  
294 increase of SPP1 as a reflection of the severity of cardiovascular<sup>35</sup>, inflammatory<sup>36</sup> and malignant<sup>37</sup>  
295 conditions. These conditions are all likely also represented in the age-matched non-motor neuron  
296 disease control group, and could explain the low relative difference in SPP1 levels between such  
297 controls and ALS patients.

298 From now on, it will be important to refine the descriptions of mechanisms in presymptomatic  
299 stages of ALS. Our data resource allows exploration of early cell-specific events and also highlights the  
300 unanticipated role of perivascular fibroblast proteins in survival prognosis. We propose that detailed  
301 studies of blood vessel-derived mechanisms should become a major focus in order improve prognostic  
302 accuracy. This could help to re-evaluate ALS etiology and to inform appropriate future therapeutic  
303 approaches although validity of SPP1 as a therapeutic target remains to be addressed with  
304 interventions. Since enlarged perivascular spaces are repeatedly observed in aging brains<sup>38</sup>, in  
305 dementia<sup>39</sup> and in other neurological disorders<sup>40</sup>, perivascular fibroblast cell activity within those  
306 spaces could represent a common therapeutic target in cerebral injury.

307

308

309

## 310 **Materials and methods**

311

### 312 ***Bioinformatics***

313 -Expression weighted cell type enrichment (EWCE)

314 Top 250 genes regulated in bulk tissue transcriptomics datasets were analyzed for overrepresentation  
315 of cell-specific transcripts derived from single cell RNA-sequencing. Single cell RNA-seq data from two

316 separate publications were used. The cortex mRNA expression data was downloaded from the  
317 associated website on the Linnarsson lab home (URL, [https://storage.googleapis.com/linnarsson-lab-  
www-blobs/blobs/cortex/expression\\_mRNA\\_17-Aug-2014.txt](https://storage.googleapis.com/linnarsson-lab-<br/>318 www-blobs/blobs/cortex/expression_mRNA_17-Aug-2014.txt))<sup>12</sup>. Duplicated gene symbols were  
319 dropped. Additional single cell RNA-seq data was obtained from the Marques et al<sup>13</sup> and downloaded  
320 from the associated webpage  
321 ([http://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE75330&format=file&file=GSE75330%5FMarqu  
es%5Fet%5Fal%5Fmol%5Fcounts2%2Etab%2Egz](http://www.ncbi.nlm.nih.gov/geo/download/?acc=GSE75330&format=file&file=GSE75330%5FMarqu<br/>322 es%5Fet%5Fal%5Fmol%5Fcounts2%2Etab%2Egz)). Cells which were annotated as belonging to the cell  
323 class "(none)" were dropped from both datasets. The two single cell datasets were joined using the  
324 `merge_two_expfiles()` function from the EWCE package<sup>41</sup>. The `level1class` for the following cells were  
325 modified: `Vsmc` to "Vascular Smooth Muscle Cell" (n=62); `Peric` to "Pericytes" (n=21); `Vend1` and  
326 `Vend2` to "Vascular Endothelial" (n=137); `astrocytes_ependymal` to "Astrocytes" (n=224); `pyramidal SS`  
327 and `CA1` to "Excitatory Neurons" (n=1338); `NFOL*`, `MFOL*` and `MOL*` to "Oligodendrocytes" (n=4528);  
328 `OPC` and `COP` to "Oligodendrocyte progenitors" (n=449); and `PPR` (Vascular and Leptomeningeal Cells)  
329 to "Perivascular Fibroblasts" (n=76). The labeling of "Microglia" (n=98) and "Interneuron" (n=290) cell  
330 categories were unchanged. Specificity data was calculated for the single cell dataset using the EWCE R  
331 package (available from [github.com/NathanSkene/EWCE](https://github.com/NathanSkene/EWCE))<sup>41</sup>.

332 A third single-cell dataset from Vanlandewijck et al.<sup>22</sup> was used for visual reference of vascular cell  
333 transcript specificity.

334  
335 The mouse *SOD1*<sup>G93A</sup> spinal cord dataset<sup>42</sup> was downloaded from GEO (accession GSE18597). Raw cell  
336 files were obtained and loaded into R using the `affy` package<sup>43</sup>. Probe annotations and mapping to  
337 HGNC symbols was done using the `biomaRt` R package<sup>44</sup>. All arrays were analyzed together, with a  
338 separate column included in the design matrix for each age and mutants at each age. Differential  
339 expression analysis was performed using the `limma` package<sup>45</sup>. The *TARDBP*<sup>Q331K/Q331K</sup> mouse dataset<sup>19</sup>  
340 RNA-seq data files were obtained from GEO (accessions GSE99353 and GSE112575). Files were  
341 combined, feature names were corrected manually. Differential expression analysis was performed  
342 using `edgeR` and `limma` packages. The human spinal cord dataset<sup>46</sup> was downloaded from GEO  
343 (accession GSE18920). The data from enriched motor neurons was dropped and only anterior horn  
344 samples were kept. Differential expression analysis was again performed using the `limma` package<sup>45</sup>  
345 controlling for gender. EWCE analysis was performed on the top 250 up/down-regulated genes, sorted  
346 based on the t-statistic, using 10,000 bootstrap replicates for each analysis.

347

348 -Gene ontology analysis

349 Human and mouse gene sets for gene ontology analysis were downloaded from [geneontology.org](http://geneontology.org).  
350 Enrichment of the gene sets within the ALS patient and mouse *SOD1*<sup>G93A</sup> expression data was analyzed  
351 with the `mroast` function from the `limma` R package, limiting for each platform the genes analyzed to  
352 the genes where data was available on the respective platform.

353

354 **ALS mouse models**

355 - *Mouse strains, housing.* The *SOD1*<sup>G93A</sup> (B6SJL-Tg(SOD1\*G93A)1Gur/J) strain used as an ALS model  
356 were a kind gift from Prof. Stefan Marklund, Tomas Brännström and Peter Andersen at Umeå  
357 University, Sweden. B6SJL-Tg(SOD1)2Gur/J mice overexpressing wild-type SOD1 were purchased from  
358 the Jackson Laboratory. Mice were housed in individually ventilated cages in a specific pathogen free  
359 facility and given free access to food and water with 12/12hour light cycle ambient room temperature  
360 19-23°C and air humidity 40-60%. Symptomatic mice were given solid drink (#95-23-100) from Nova  
361 SCB. Transgene-bearing mice were identified by PCR genotyping as described previously<sup>7</sup>. *SOD1* copy  
362 number was determined with qPCR using fluorescent probes for *hSOD1* and *mApoL* and was carried as  
363 described in guidelines by the Jackson Laboratory<sup>47</sup>. Mice with more than a 0.5 dCt differences from  
364 the *SOD1*<sup>G93A</sup> reference DNA (Jackson Laboratories) were discarded from the colony.

365

366 **Ethical approval**

367 All the work involving animal or human subjects or tissues has been carried out in accordance with the  
368 Code of Ethics of the World Medical Association (Declaration of Helsinki) and with national legislation  
369 as well as our institutional guidelines. Animal experiments were approved and performed according to  
370 the guidelines of the North Stockholm Animal Ethics Committee. ALS patient plasma or CSF collection  
371 was approved by the ethics committees at Ulm University (application number 20/10), UMC Utrecht  
372 (SL/nb/ 16/004075) Leuven University (ML4073) and Karolinska Institute (2018/1605).

373

374 **Immunostaining**

375 -ALS Subjects

376 Post-mortem material (Table 1.) <https://figshare.com/s/4c89e57b39620d020f8d> was obtained at  
377 autopsy from ALS patients at the department of (Neuro)pathology of the Amsterdam UMC, Academic  
378 Medical Center, University of Amsterdam, the Netherlands. All patients fulfilled the diagnostic criteria  
379 for ALS (El Escorial criteria)<sup>48</sup> as reviewed independently by two neuropathologists. All patients with  
380 ALS died from respiratory failure. Control spinal cord tissue was obtained from patients who had died  
381 from a non-neurological disease. Both ALS and control patients included in the study displayed no  
382 signs of infection before death. Informed consent was obtained for the use of brain tissue and for  
383 access to medical records for research purposes and approval was obtained from the relevant local  
384 ethical committees for medical research. All autopsies were performed within 12 hours after death.

385

386 -Tissue preparation

387 Paraffin-embedded tissue was sectioned at 6 µm and mounted on pre-coated glass slides (StarFrost,  
388 Waldemar Knittel Glasbearbeitungs GmbH, Braunschweig, Germany). Representative sections of all  
389 specimens were processed for haematoxylin and eosin and Klüver-Barrera.

390

391 -Immunohistochemistry on spinal cord samples from sporadic ALS patients

392 For immunohistochemistry on human spinal cord samples, formalin-fixed paraffin-embedded 6  $\mu$ m  
393 thick sections were deparaffinized in xylene and rinsed in graded ethanol (100%, 95%, 70%). Antigen  
394 retrieval was performed in 0.01M HCl with 0,5% Pepsin (Sigma Aldrich, Darmstadt, Germany, #P7012)  
395 at 37°C for 15 minutes in a water bath followed by incubation with a given primary antibody (Collagen  
396 IV, MS-747-S, Thermo Sci, 1:50; COL6A1, HPA019142, Atlas antibodies, 1:500; SPP1, AF1433, R&D,  
397 1:50; LAMA1, Sigma L9393, 1:50). Incubation with primary antibody was performed overnight at 4°C .  
398 After washing in PBS, sections were stained with a polymer based peroxidase immunohistochemistry  
399 detection kit (Brightvision plus kit, ImmunoLogic, Duiven, the Netherlands) according to the  
400 manufacturer's instructions. Staining was performed using Bright DAB substrate solution  
401 (ImmunoLogic, Duiven, the Netherlands). Sections were dehydrated in alcohol and xylene and  
402 coverslipped.

403 For double immunohistochemistry, sections were incubated with Brightvision poly-alkaline  
404 phosphatase-goat-anti mouse (Immunologic, Duiven, The Netherlands) for 30 min at RT and washed  
405 with PBS. Sections were washed with Tris-HCl buffer (0.1 M, pH 8.2) to adjust the pH. Alkaline  
406 phosphatase activity was visualized with the alkaline phosphatase substrate kit III Vector Blue (SK-  
407 5300, Vector laboratories Inc., CA, USA). The first primary antibody was cooked off by cooking in  
408 10mM citrate buffer pH6.0 for 10 min at 100°C in a pressure cooker and subsequently incubated with  
409 the second primary antibody overnight at 4°C . The next day the sections were incubated with rabbit-  
410 anti-goat (SouthernBiotech #SBA 6164-01) for SPP1 for 15 min and then with Brightvision goat-anti-  
411 rabbit poly HRP for 30 min at RT, and washed with PBS. Horseradish peroxidase was visualized with  
412 filtered 5x10<sup>-4</sup> % w/v AEC in in 0.05M acetate buffer pH 4,9 and 1x10<sup>-4</sup> % H<sub>2</sub>O<sub>2</sub>. Sections incubated  
413 without primary antibodies were blank.

414

415 Sporadic ALS was defined as without the presence of C9ORF72 hexanucleotide repeat expansion or  
416 mutations analyzed by targeted NGS analysis. The NGS panel consists of TARBP, ALS2, ErbB4, NEK1,  
417 MATR3, VCP, SIGMAR1, c9orf72, c19orf12, OPTN, HNRNPA1, DAO, SPG11, FUS, GRN, PNPLA6, SOD1,  
418 CHCHD10, NEFH and UBQLN2.

419 For histopathological scoring, all labeled tissue sections were evaluated by two independent observers  
420 blinded to clinical data for the presence or absence of various histopathological parameters and  
421 specific immunoreactivity (IR) for the different markers. Hematoxylin-Eosin (HE) and Nissl stained  
422 slides were used to evaluate the neuronal and glial components of the tissues. The intensity of GFAP  
423 and HLA-DR (MHC-II) immunoreactive staining was evaluated using a scale of 0-3 (0: -, no; 1: +/-, weak;  
424 2: +, moderate; 3: ++, strong staining). The frequency of GFAP and HLA-DR positive cells [(1) rare; (2)  
425 sparse; (3) high] was also evaluated to give information about the relative number of positive cells  
426 within the spinal cord (lumbal region). As proposed before<sup>49,50</sup>, the product of these two values

427 (intensity and frequency scores) was taken to give the overall score (total score; immunoreactivity  
428 score; IRS). **Key scoring: Frequency:** (1) < 1-10 % (2) 11-50 % moderate; (3) > 50 %. **Intensity:** 0: not  
429 present; 1+, weak; 2+, moderate; 3+, strong.

430 Antibodies used for scoring included: Glial fibrillary acidic protein (GFAP; polyclonal rabbit, Z0334,  
431 DAKO, Glostrup, Denmark; 1:4000), neuronal nuclear protein (NeuN; mouse clone A60, MAB377  
432 Merck-Millipore), major histocompatibility complex (MHC) class II antigen (HLA)-DP, DQ, DR (mouse  
433 clone CR3/43; M0775, DAKO, Glostrup, Denmark; 1:400) and CD68 (mouse clone PG-M1, M0876,  
434 DAKO; 1:200).

435

436 -Mouse tissue immunofluorescence and confocal microscopy

437 Following PBS and PFA perfusion, mouse tissues were incubated with 30% sucrose solution overnight,  
438 embedded in frozen section media and snap frozen on dry ice. 16 µm sections were cut on a Micron  
439 cryostat for histology staining and quantifications. The following antibodies were used for  
440 immunostainings: Podocalyxin (AF1556 R&D; 1:250), Collagen IV (2150-1470 Serotec; 1:250), SPP1  
441 (HPA027541, Atlas antibodies 1:250), COL6A1 (HPA019142, Atlas antibodies, 1:250) and LAMA1  
442 (L9393, Sigma, 1:500). Secondary goat antibodies conjugated to Alexa Fluor 488, 555, 594 or 647 were  
443 purchased from Life Technologies and used at 1:500. Images were acquired using Zeiss LSM 700  
444 confocal microscope and Zen software.

445

#### 446 ***Transmission electron microscopy***

447 Transmission electron microscopy was performed according to standard protocols. Briefly, mice were  
448 anesthetized and perfused with PBS. Spinal cords were excised and fixed in the fixation solution buffer  
449 (2% glutaraldehyde, 0.5% paraformaldehyde, 0.1 M cacodylate, 0.1 M sucrose, 3 mM CaCl<sub>2</sub>) and  
450 washed in 0.1 M cacodylate buffer (pH, 7.4) before staining in 2% OsO<sub>4</sub> in cacodylate buffer for 1 hour  
451 at room temperature. Samples were dehydrated and *en bloc* staining was performed in 2% uranyl  
452 acetate in absolute ethanol for 1 hour at room temperature; samples were then taken through an  
453 Epon 812/acetone series and embedded at 60°C in pure Epon 812. Thin sections of 70 nm thickness  
454 were made on a Leica EM UC6 Ultratome and mounted on Formvar-coated copper slot grids. Post  
455 staining was done with 2% aqueous acetate (pH, 3.5) and Venable and Coggeshall lead citrate. Grids  
456 were analyzed on an FEI TECNAI electron microscope.

457

#### 458 ***Image analysis***

459 Quantifications of DAB and two-colour staining area from were performed with IHC image analysis  
460 toolbox for ImageJ. Two full frame 4x pictures adjusted for white background from each patient spinal  
461 cord tissue were analyzed to get per-patient average n result. Quantification for immunofluorescence  
462 in mouse spinal cords were performed by setting a common pixel intensity threshold and calculating  
463 pixel-intensity within positive area using ImageJ. A typical 3-4 frames taken at 20x objective were used

464 to obtain per mouse average intensity. Pictures are presented as 3D renderings of 16µm thick sections  
465 from 10 z stacks and presented on white background using open source Icy software ver. 1.9.9.1.  
466 Perivascular space area in electron microscopy images was calculated from 3-5 regions of interest  
467 within the ventral horn grey matter that contained blood vessels acquired once from a group of 4 mice  
468 per genotype. First, we measured the areas outlined by mural and astrocyte basement membrane  
469 (BM). The gap was then calculated by subtracting the mural from astrocyte BM areas and divided by  
470 the mural BM area to normalize for individual vessel diameter.

471

#### 472 ***Statistics and reproducibility***

473 For immunostaining and transmission electron microscopy quantifications blinding was performed by  
474 third party concealment of treatments or genotypes and assignment of numeric codes to each group.  
475 Experiments were performed twice unless indicated otherwise. Representative graphs such as  
476 basement membrane structure presented by transmission electron microscopy or by histology  
477 stainings were based on observations found in 4 out of 4 analyzed animals or human cases or  
478 quantified in respective data panels. Data are expressed as median ± SEM. Box plots represent median,  
479 Statistical comparisons were made with GraphPad Prism or R analysis software using Student's or  
480 Wilcoxon tests (2-tailed unless otherwise indicated). For all tests, a p-value of 0.05 or less was  
481 considered significant.

482

#### 483 ***Patient cohorts for proteomics plasma profiling.***

484 Netherlands (Utrecht). Plasma samples and patient selection from Utrecht UMC was performed as  
485 described previously<sup>51</sup>. In short: patients diagnosed with suspected, possible, probable or definite ALS  
486 according to the El Escorial criteria were included. Patients with progressive muscular atrophy, primary  
487 lateral sclerosis and progressive bulbar palsy were excluded from analysis. In order to determine  
488 whether a patient fulfilled the El Escorial criteria, the correspondence of the neurologist, including  
489 results of neurophysiological examination, was scrutinized. The controls included plasma from general  
490 hospital admission from patients without motor neuron disease.

491 Germany (Ulm). Patients were diagnosed according to the El Escorial criteria including possible ALS.

492 The controls were patients were sampled because of headache, differential diagnoses subarachnoid  
493 hemorrhage or differential diagnoses meningitis.

494 Belgium (Leuven). A total of 134 patients with ALS and 27 control patients were consecutively included  
495 in a prospective manner between April 2014 and September 2016. All patients were seen at the  
496 Neuromuscular Reference Center (NMRC) of the University Hospitals Leuven. The patients with ALS  
497 were diagnosed according to the Awaji and revised El Escorial criteria. Sampling occurred during the  
498 diagnostic phase of the patient with ALS. The control cohort consisted out of 20 neurologic non-motor  
499 neuron disease controls and 7 patients with nonspecific subjective complaints for whom an underlying  
500 neurologic condition was ruled out upon neurologic examination.

501 Blood samples were obtained during the first visit to the NMRC. Serum was extracted after 10 minutes  
502 of centrifugation at 1955g, transferred into coded cryovials of 1mL and stored at -80°C.  
503 Sweden (Stockholm). Recruited patients received a diagnosis of ALS the Karolinska ALS center, an  
504 outpatient clinic that manages all ALS patients in the Greater Stockholm area. ALS patients who met  
505 the revised El Escorial criteria for definite, probable and probable laboratory-supported ALS were  
506 included.<sup>52</sup> Patients that emigrated from the Stockholm region during the study period were excluded.  
507 ALS patients were regularly evaluated by a neurologist, who registered their clinical characteristics in  
508 the Swedish Motor Neuron Disease Quality Register.<sup>53</sup> We further included siblings and partners of the  
509 ALS patients as healthy controls. We performed a neurological exam of the controls but did not assess  
510 their health status further. Healthy controls were recruited shortly after the diagnosis of the index  
511 patient (usually within six months). Of the included siblings, one sibling was related to an ALS patient  
512 that carried a *C9orf72* repeat expansion, and one sibling was related to an ALS patient with a *SOD1*  
513 mutation. For ALS patients, CSF and blood were collected at the time of diagnosis (+/- 90 days). All  
514 patients were offered repeated blood (once every three to six months) and CSF sampling (once per  
515 year). CSF and blood were collected from the ALS mimics during the diagnostic work-up, and from the  
516 healthy controls shortly after the diagnosis of the index patient. CSF was obtained through lumbar  
517 puncture directly into polypropylene tubes. The CSF samples were centrifugated for ten minutes at  
518 400 g at room temperature. Plasma was collected close to the date of the lumbar puncture, and  
519 centrifugated for ten minute at 2,000 g at room temperature. Aliquots of 1 ml (CSF) and 800  
520 µL (plasma) were directly frozen and stored in -80°C. NFL content in CSF was measured using the  
521 UmanDiagnostics ELISA kit, Sweden (cat no 10-7001).

522

### 523 ***Affinity proteomics of plasma***

524 -Antibody selection

525 Protein targets for plasma analysis were chosen based on mRNA transcript specificity to perivascular  
526 fibroblasts, their induction in the human or mouse transcriptomics and based on antibody availability  
527 within the Human Protein Atlas (HPA) project. A total of 32 antigen-purified and protein microarray  
528 validated antibodies were selected for 15 unique proteins. Results presented in the figures were based  
529 on antibodies against COL6A1 (HPA019142), SPP1 (HPA005562 or HPA027541).

530

531 -Suspension bead array

532 The procedure for suspension bead arrays was performed as described previously<sup>54,55</sup>. In short,  
533 samples were distributed in 96-well microtiter plates, diluted 1:10 in phosphate buffered saline, and  
534 the protein content directly labeled with biotin. For the bead array, antibodies were immobilized onto  
535 magnetic color-coded beads with one bead identity corresponding to a certain antibody. Samples were  
536 then further diluted 1:50 in an assay buffer, heat treated at 56°C for 30 min, combined into a 384-well  
537 microtiter plate, and incubated with the bead array at RT on a shaker overnight. Unbound proteins

538 were removed by washing and proteins on the beads were detected through a streptavidin-conjugated  
539 fluorophore (Invitrogen.com). Results from the FlexMap3D instrument (Luminex Corp.Com) were  
540 reported per bead identity as median fluorescence intensities (MFI). The CSF analysis was performed  
541 similar as for plasma with some minor adjustments as reported previously in (Pin et al 2019 (PMID  
542 31432421). First, 15 ul of the samples were labelled with biotin in a protein containing buffer (PBS  
543 supplemented with 0.5% (w/v) bovine serum albumin and 0.1% (w/v) rabbit IgG at an end dilution of  
544 1/2. For the assay, the samples were diluted 1/8 in assay buffer to and end volume of 50 ul before heat  
545 treatment.

546

547 -Antibody validation

548 The SPP1 antibody was validated with Western blot. Plasma samples were diluted 1:40 in MilliQ and  
549 proteins separated on a gel before blotted on a membrane (all Invitrogen). Detection antibody was  
550 applied at 1 µg/mL and binding allowed at 4°C overnight followed by readout through a  
551 chemiluminescent substrate (BioRad.Com).

552

553 -Suspension bead array data processing and statistical analysis

554 Data were processed and visualized in R (v. 3.6.1). Samples with less than 20 counted beads per  
555 identity as well as technically failed samples were excluded from further analysis. To diminish any plate  
556 effect, multidimensional normalization was applied with the assumption that the mean of each plate  
557 for each antibody should be close. MA-LOESS normalization was thereafter applied to reduce effects  
558 associated to the different assay plates followed by log transformation of the datasets. Coefficients of  
559 variation were calculated based on a pool of samples analyzed in triplicate in each 96-well plate. For  
560 combined cohort analysis the datasets were median scaled. Association between groups was assessed  
561 by Wilcoxon rank sum tests.

562

563 -Optimization of plasma value cut-offs

564 To consider the effect of elevated protein levels, we determined optimal cut points to divide cohorts in  
565 “low” or “high” level groups for the proteins in question. Cut points were based on patient protein  
566 fluorescence intensity values in respect to survival. We used maximally selected rank statistics with the  
567 logrank test similar to Kaplan-Meier analysis using the “maxstat” package v.0.7-25 ([https://CRAN.R-  
568 project.org/package=maxstat](https://CRAN.R-project.org/package=maxstat)). All patient protein values were tested and considered as cut points  
569 excluding 10 percent from the upper and lower end of the protein values range. Survival was analyzed  
570 using “A Package for Survival Analysis in S” version 2.38. <https://CRAN.R-project.org/package=survival>.  
571 Kaplan-Meier graphs were plotted based on the empirical threshold groups using “survminer” package  
572 v. 0.4.4 (<https://CRAN.R-project.org/package=survminer>). The protein values for COL6A1 in the Belgian  
573 and Swedish cohorts yielded no clear cut point for survival groups and results shown are based on the  
574 threshold at the median value.



575

576 -Cox proportional hazard

577 In order to assess the risk of having elevated protein levels in comparison to other variables we  
578 created univariate and multivariate Cox proportional hazard models. Variables tested were disease  
579 onset type, gender and age at sampling. Both univariate and multivariate models were corrected for  
580 sampling delay and cohort effect whenever applicable. Sampling delay was added as a stratification  
581 term, whereas we used a cluster term on cohort identities to adjust for possible unknown correlations  
582 within cohorts. The latter being a generalized estimating equation (GEE) term. Above analyses were  
583 performed in R version 3.6.1 (<https://www.R-project.org/>). Patient plasma data and analysis scripts for  
584 cut-off optimization and Cox proportional hazard are available at:

585 [https://github.com/lewandowski/PVF\\_Manuscript](https://github.com/lewandowski/PVF_Manuscript)

586

### 587 **Analysis of ALSFRS data at the time of plasma/CSF sampling**

588 In this analysis, we aimed to assess the ALSFRS value at the time point of plasma/CSF sampling. Table 2  
589 <https://figshare.com/s/4c89e57b39620d020f8d> displays the numbers of ALS patients who had ALSFRS  
590 data available in each of the cohorts, and who were included in the analysis. Patients were included if  
591 their date of clinical onset was known (at which point the ALSFRS = 48) and they had at least one  
592 ALSFRS measurement available with known date.

593 For each cohort, a subset of patients had ALSFRS available on the same date as the plasma/CSF  
594 sampling or within  $\pm 10$  days. For these patients, we used the actual ALSFRS measure.

595

596 For the remaining patients, we estimated the ALSFRS at the plasma/CSF sampling time point by fitting  
597 the available ALSFRS experimental data with a model. When only two data points were available  
598 (clinical onset at ALSFRS = 48 and one more ALSFRS), we performed a linear fit through the two points,  
599 and used that formula to estimate ALSFRS at the plasma/CSF sampling time point. When we had three  
600 or more ALSFRS data points, we used a sigmoid curve, following a method recently published<sup>56</sup>. This  
601 study has determined that when patients are followed longitudinally over multiple time points, the  
602 ALSFRS data follows a sigmoidal curve, being characterized by an initial phase of relatively gradual  
603 decline near the initial point of disease onset, followed by a time interval of relatively fast decline, and  
604 leading to a final stage of slower decline when patients are at a severe stage. The sigmoid function  
605 allows describing the whole ALSFRS temporal progression using two parameters, D50 and dx, as  
606 expressed by the following formula<sup>56</sup>:

607

$$y = \frac{48}{1 + e^{\frac{(x-D50)}{dx}}}$$

608

609 where:

610  $y$  = individual ALSFRS scores over time for a given patient  
611  $x$  = time point, number of months after clinical onset  
612 numerator = corresponds to the ALSFRS score (48) assumed at the time of clinical onset  
613 D50 = parameter that is equal to the number of months after clinical onset ( $x$ ) when the ALSFRS  
614 score ( $y$ ) drops to half of its value at clinical onset, i.e., when  $y$  drops to 24  
615  $dx$  = slope of the sigmoid curve at  $x = D50$ , with low  $dx$  values corresponding to steeper decline,  
616 and high  $dx$  values corresponding to slower decline at  $x = D50$

617

618 For each individual patient, the ALSFRS-R scores were fitted to the sigmoid function using Microsoft®  
619 Excel Add-In Solver tool, by an iterative procedure used to simultaneously estimate the D50 and  $dx$   
620 values that provide the best fit for the longitudinal ALSFRS-R data. The sigmoid model was applied if  
621 there were at least 3 time points for each patient.

622

### 623 **Data availability**

624 Human ALS patient transcriptome datasets and analysis scripts are available at:

625 [https://github.com/NathanSkene/ALS\\_Human\\_EWCE](https://github.com/NathanSkene/ALS_Human_EWCE)

626 *SOD1*<sup>G93A</sup> mouse transcriptome datasets and analysis scripts are available at:

627 [https://github.com/NathanSkene/ALS\\_Mouse\\_EWCE](https://github.com/NathanSkene/ALS_Mouse_EWCE)

628 *TARDBP*<sup>Q331K/Q331K</sup> mouse transcriptome datasets and analysis scripts are available at:

629 [https://github.com/szczepinskaa/ALS\\_TDP-43.git](https://github.com/szczepinskaa/ALS_TDP-43.git)

630 Images and analysis scripts for human histology are deposited and available under the following links:

631 SPP1 in Red - <https://figshare.com/s/3ad17913ca4fb2e99b80>

632 COL6A1 in Red - <https://figshare.com/s/f318c332f4fbc31b48a0>

633 SPP1 in DAB - <https://figshare.com/s/ee294c7715f77db7df14>

634 COL6A1 in DAB - <https://figshare.com/s/8c0ecbe9b3fbc63cfef0>

635 Additional data tables are available at: <https://figshare.com/s/4c89e57b39620d020f8d>

636 Patient plasma data and analysis scripts for cut-off optimization and Cox proportional hazard are  
637 available at: [https://github.com/lewandowskilab/PVF\\_Manuscript](https://github.com/lewandowskilab/PVF_Manuscript)

638

### 639 **Code availability**

640 Human ALS patient transcriptome analysis scripts are available at:

641 [https://github.com/NathanSkene/ALS\\_Human\\_EWCE](https://github.com/NathanSkene/ALS_Human_EWCE)

642 *SOD1*<sup>G93A</sup> mouse transcriptome analysis scripts are available at:

643 [https://github.com/NathanSkene/ALS\\_Mouse\\_EWCE](https://github.com/NathanSkene/ALS_Mouse_EWCE)

644 *TARDBP*<sup>Q331K/Q331K</sup> mouse transcriptome analysis scripts are available at:

645 [https://github.com/szczepinskaa/ALS\\_TDP-43.git](https://github.com/szczepinskaa/ALS_TDP-43.git)

646 Images and analysis scripts for human histology are deposited and available under the following links:

647 SPP1 in Red training model - <https://figshare.com/s/3ad17913ca4fb2e99b80>  
648 COL6A1 in Red training model - <https://figshare.com/s/f318c332f4fbc31b48a0>  
649 SPP1 in DAB training model - <https://figshare.com/s/ee294c7715f77db7df14>  
650 COL6A1 in DAB training model and macro - <https://figshare.com/s/8c0ecbe9b3fbc63cfef0>  
651 Patient plasma analysis scripts for cut-off optimization and Cox proportional hazard are available at:  
652 [https://github.com/lewandowskilab/PVF\\_Manuscript](https://github.com/lewandowskilab/PVF_Manuscript)

653

654

655

656

657

#### 658 **Extended data figures**

659 (see separate document)

660

#### 661 **Acknowledgements**

662 S.A.L. is supported by the Olle Engkvist Byggmästare Foundation (SLS-499431), Ulla-Carin Lindquists  
663 stiftelse för ALS-forskning, Åhléns foundation (mA2/h17, 203074), Thierry Latran Foundation (FIB-ALS)  
664 and Neuroförbundet. N.S. was supported by Wellcome Trust (108726/Z/15/Z), Edmond J. Safra  
665 Foundation, Lily Safra and UK Dementia Research Institute. We thank the ALS Stichting grant "The  
666 Dutch ALS Tissue Bank" (E.A.) and Netherlands Brain Bank (E.H.) for providing the ALS tissue samples.  
667 We acknowledge the team who helped in the collection of ALS tissue samples (Prof. dr. D. Troost,  
668 Prof. dr. M. de Visser, Dr. A.J. van der Kooi and Dr. J. Raaphorst). U.K. and C.I. are supported by  
669 Björklunds Fund, the Ulla-Carin Lindquist Foundation, Neuro Sweden, SLL Halsä Medicin och Teknik.  
670 E.R-V. is supported by Swedish Alzheimer Foundation (Alzheimerfonden), Swedish Dementia  
671 Association (Demensfonden), Gun & Bertil Stohne's Foundation, and Gamla Tjänarinnor Foundation.  
672 This project has received funding from the European Research Council (ERC) under the European  
673 Union's Horizon 2020 research and innovation program (grant agreement n° 772376 – ESCORIAL  
674 awarded to J.V. R.A.H. is supported by Alltid Litt Sterkere, AlzheimerFonden, Swedish Medical  
675 Research Council, Swedish brain foundation and Karolinska Institutet. M.L. was supported by a grant  
676 from the Knut and Alice Wallenberg Foundation (2012.0091). This study was also supported by grants  
677 to P.N. from the Swedish FTD initiative funded by the Schörling Family Foundation and the KTH Center  
678 for Applied Precision Medicine (KCAP) funded by the Erling-Persson Family Foundation. P.V.D. holds a  
679 senior clinical investigatorship of FWO-Vlaanderen and is supported by E. von Behring Chair for  
680 Neuromuscular and Neurodegenerative Disorders, the ALS Liga België and the KU Leuven funds "Een  
681 Hart voor ALS", "Laeversfonds voor ALS Onderzoek" and the "Valéry Perrier Race against ALS Fund".  
682 Several authors of this publication are member of the European Reference Network for Rare  
683 Neuromuscular Diseases (ERN-NMD). We would also like to thank Kjell Hultenby at Karolinska

684 Institutet EM facility, Gudný Ella Thorlacius, Mun-Gwan Hong Sofia Bergström, Jamil Yousef, Heela  
685 Sarlus and Ali Manouchehrinia for support with data analysis and Tomas Brännström with Matthew  
686 Marklund for assistance with selection and harvest of *SOD1<sup>G93A</sup>* mouse tissues. S.A.L would like to  
687 thank the unpaid interns on Erasmus scholarships for their efforts and contributions.

688

#### 689 **Author contributions**

690 N.S. and S.A.L. designed transcriptomics enrichment experiments.

691 N.S. and A.S. performed the computational analysis of enrichment experiments.

692 M.T., I.S.A., A.D. and P.L. performed additional transcriptomics analysis.

693 M.L., J.D.G. and S.A.L. facilitated and performed mouse histology and immunostaining.

694 I.V.G.A. performed histology validation experiments.

695 L.E., M.T. and S.A.L. performed electron microscopy imaging and quantifications.

696 R.A.H., E.H., J.A., C.M. and E.A., facilitated and performed human histology staining.

697 Clinical centers in: Ulm (A.H., A.L.), Utrecht (H.V.B., J.V.), Leuven (J.D.V., M.D.S., K.P., P.V.D.) and  
698 Stockholm (C.I., U.K.) - designed patient and control cohorts and provided plasma or CSF samples.

699 S.A.L., A.M. and J.R. designed and chose protein targets for plasma profiling.

700 A.M., J.R., J.O., F.S., S.W. M.U., P.N., E.R-W. and C.I. facilitated and performed plasma profiling, survival  
701 associations, and clinical parameter statistics.

702 S.A.L. wrote the manuscript with input from the co-authors.

703 P.N., and S.A.L. oversaw all aspects of the study.

704

705 **Competing Interests.** We declare that the authors do not have competing interests.

706

707

#### 708 **Figure Legends** (for main text figures)

709

710 **Figure 1. ALS patients show increased transcriptional activity of perivascular fibroblast cell gene**  
711 **markers which occur at presymptomatic disease stage in *SOD1<sup>G93A</sup>* and *TARDBP<sup>Q331K</sup>* mice. (A)**

712 Schematic of expression weighted cell type enrichment (EWCE) analysis. Cell type-specific gene  
713 rankings from single cell sequencing data allow to infer cell type activity in bulk tissue transcriptomes.

714 **(B)** Enrichment z - scores for up and down-regulated genes in ten cell type classes in sALS patients,  
715 *SOD1<sup>G93A</sup>* and *TARDBP<sup>Q331K</sup>* mice. In the *SOD1<sup>G93A</sup>* mice, onset of neuromuscular junction decoupling (8  
716 weeks) and clinical symptoms (peak body weight - 16 weeks) are indicated with arrowheads. P-values  
717 are specified in Ext. data Fig. 2. **(C)** Cell type RNA specificity for genes enriched in perivascular

718 fibroblasts. N number of cells per category is described in Ext. data Fig. 1. Bars show median count of  
719 RNA per cell  $\pm$  SEM. **(D)** Perivascular fibroblast specific gene activity in sALS patient spinal cords. sALS  
720 n=12, Ctrl n=8. Boxplots show median, 2nd-3rd quartile and whiskers show  $\pm$ 1.5 of the IQR. Two-

721 tailed t-test p-values are indicated next to brackets. (E) Expression of perivascular fibroblast enriched  
722 genes in *SOD1<sup>G93A</sup>* (p-value for ANOVA in early (4-10 weeks) and late (14-18 weeks) timepoints) and  
723 *TARDBP<sup>Q331K</sup>* mice (p-values for two-tailed t-test). (F) Perivascular fibroblasts have distinct mRNA  
724 expression markers (median RNA molecule count with  $\pm$ SEM). (G) Schematic illustration represents  
725 reported PVF location between astrocytes and mural/endothelial cells. (H) Transmission electron  
726 microscopy of mouse spinal cord tissue points to location of perivascular fibroblasts cells (PVF - shaded  
727 red) between basement membrane layers that delineate astrocyte endfeet (AC - blue) and mural cells  
728 (MC - brown), (EC - yellow: endothelial cell, aBM - astrocyte basement membrane, mBM - mural  
729 basement membrane). Scale bar 2 $\mu$ m.

730

731 **Figure 2. Perivascular fibroblast marker proteins COL6A1 and SPP1 accumulate in enlarged**  
732 **perivascular spaces during ALS progression.** (A) Col6a1 and Spp1 mRNA specificity within CNS cell  
733 types. Bars represent relative count of RNA per cell  $\pm$  SEM. (B) COL6A1 and SPP1 histochemistry in sALS  
734 and control spinal cords, bar: 10 $\mu$ m. (C) Quantifications of human tissue histochemistry from full frame  
735 4x photos. sALS and Ctrl n=4 individuals (2-tailed t-test p-value). All boxplots show median, 2nd-3rd  
736 quartile and whiskers show  $\pm$ 1.5 of the IQR. (D) Col6a1 and Spp1 accumulate around blood vessels  
737 (outlined with podocalyxin - cyan) in 14 week *SOD1<sup>G93A</sup>* mouse spinal cords. Immunofluorescence z-  
738 stack renderings of 16 $\mu$ m thick sections, bars: 100 $\mu$ m (overview), 10 $\mu$ m (insert). (E) Quantifications of  
739 immunofluorescence stainings in mice from full-frame 20x photos. *SOD1<sup>G93A</sup>* and BL/6 n=4, *SOD1<sup>wt</sup>* n=3  
740 mice (2-tailed t-test p-value). (F) Increased perivascular spaces appear in presymptomatic (8 weeks)  
741 *SOD1<sup>G93A</sup>* mice spinal cords. Immunofluorescence for vascular (Col4a1) and astrocyte (Lama1)  
742 basement membranes, bar: 10 $\mu$ m. (G) Electron microscopy (tEM) of increased perivascular spaces in  
743 14 week *SOD1<sup>G93A</sup>* mice. Astrocyte (red) and vascular (blue) basement membranes are indicated with  
744 lines. Perivascular space is indicated with asterisk, bar: 5 $\mu$ m. (H) COL6A1 and SPP1 accumulate within  
745 increased perivascular spaces (outlined with COL4A1) in spinal cords of sALS patients. 2 color  
746 histochemistry, bar: 10 $\mu$ m. (I) Quantifications of COL6A1 and SPP1 immunostainings from panel H.  
747 sALS and Ctrl n=4 individuals (two-tailed t-test p-value). (J) Quantifications of perivascular space  
748 increase in *SOD1<sup>G93A</sup>* mice from panel G (*SOD1<sup>G93A</sup>* and BL/6 n=4 mice) and in sALS patients from panel  
749 H (sALS and Ctrl n=4 individuals) (two-tailed t-test p-value). (K) Schematic representation of  
750 perivascular fibroblast activity and enlarged perivascular spaces in ALS.

751

752 **Figure 3. Prognostic value of SPP1 protein in plasma of ALS patients.** (A) Relative levels of SPP1  
753 protein in plasma as measured by the HPA027541 antibody. Threshold selection and Kaplan-Meier  
754 survival estimates of ALS patients in discovery cohorts (Netherlands, Germany and Belgium, n=452)  
755 and in the replication cohort (Sweden, n=122). Red color indicates thresholded protein level.  
756 Thresholds are established using maximally selected log rank statistics. Boxplots show median, 2nd-3rd  
757 quartile and whiskers show  $\pm$ 1.5 of the IQR. Survival probability graphs show proportion of censored

758 patients within each arm and Kaplan-Meier logrank p-values. **(B-C)** Uni- and multivariate Cox  
759 proportional hazard models for continuous increase of plasma SPP1 relative to hazard ratios indicated  
760 by bulbar onset type, neurofilament light (NFL) in CSF (in the replication cohort), gender and plasma  
761 sampling age. Whiskers represent 95% CI. Cohort identity was additionally used as covariate in  
762 multivariate models.  
763

764 **References (to main text and materials and methods)**

- 765 1. Cook, C. & Petrucelli, L. Genetic Convergence Brings Clarity to the Enigmatic Red Line in ALS.  
766 *Neuron* **101**, 1057–1069 (2019).
- 767 2. Spiller, K. J. *et al.* Microglia-mediated recovery from ALS-relevant motor neuron degeneration  
768 in a mouse model of TDP-43 proteinopathy. *Nat. Neurosci.* **21**, 329–340 (2018).
- 769 3. Boillée, S. *et al.* Onset and progression in inherited ALS determined by motor neurons and  
770 microglia. *Science* **312**, 1389–1392 (2006).
- 771 4. Lewandowski, S. A. *et al.* Presymptomatic activation of the PDGF-CC pathway accelerates onset  
772 of ALS neurodegeneration. *Acta Neuropathol.* **131**, 453–64 (2016).
- 773 5. Brown, R. H. & Al-Chalabi, A. Amyotrophic Lateral Sclerosis. *N. Engl. J. Med.* **377**, 162–172  
774 (2017).
- 775 6. Al-Chalabi, A. *et al.* Analysis of amyotrophic lateral sclerosis as a multistep process: A  
776 population-based modelling study. *Lancet Neurol.* **13**, 1108–1113 (2014).
- 777 7. Rosen, D. R. *et al.* Mutations in Cu/Zn superoxide dismutase gene are associated with familial  
778 amyotrophic lateral sclerosis. *Nature* **362**, 59–62 (1993).
- 779 8. Sreedharan, J. *et al.* TDP-43 mutations in familial and sporadic amyotrophic lateral sclerosis.  
780 *Science* **319**, 1668–72 (2008).
- 781 9. Wu, C.-H. *et al.* Mutations in the profilin 1 gene cause familial amyotrophic lateral sclerosis.  
782 *Nature* **488**, 499–503 (2012).
- 783 10. Kang, S. H. *et al.* Degeneration and impaired regeneration of gray matter oligodendrocytes in  
784 amyotrophic lateral sclerosis. *Nat. Neurosci.* **16**, 571–9 (2013).
- 785 11. Zhong, Z. *et al.* ALS-causing SOD1 mutants generate vascular changes prior to motor neuron  
786 degeneration. *Nat. Neurosci.* **11**, 420–422 (2008).
- 787 12. Zeisel, A. *et al.* Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-  
788 seq. *Science (80-. )*. **347**, 1138–42 (2015).
- 789 13. Marques, S. *et al.* Oligodendrocyte heterogeneity in the mouse juvenile and adult central  
790 nervous system. *Science (80-. )*. **352**, 1326–1329 (2016).
- 791 14. Zeisel, A. *et al.* Molecular Architecture of the Mouse Nervous System. *Cell* **174**, 999–1014.e22  
792 (2018).
- 793 15. Skene, N. G. & Grant, S. G. N. Identification of vulnerable cell types in major brain disorders  
794 using single cell transcriptomes and expression weighted cell type enrichment. *Front. Neurosci.*  
795 **10**, 1–11 (2016).
- 796 16. Clark, J. A., Southam, K. A., Blizzard, C. A., King, A. E. & Dickson, T. C. Axonal degeneration, distal  
797 collateral branching and neuromuscular junction architecture alterations occur prior to  
798 symptom onset in the SOD1G93A mouse model of amyotrophic lateral sclerosis. *J. Chem.*  
799 *Neuroanat.* **76**, 35–47 (2016).
- 800 17. Hall, E. D., Oostveen, J. A. & Gurney, M. E. Relationship of microglial and astrocytic activation to

- 801 disease onset and progression in a transgenic model of familial ALS. *Glia* **23**, 249–56 (1998).
- 802 18. Weydt, P., Hong, S. Y., Kliot, M. & Möller, T. Assessing disease onset and progression in the  
803 SOD1 mouse model of ALS. *Neuroreport* **14**, 1051–4 (2003).
- 804 19. White, M. A. *et al.* TDP-43 gains function due to perturbed autoregulation in a Tardbp knock-in  
805 mouse model of ALS-FTD. *Nat. Neurosci.* **21**, 552–563 (2018).
- 806 20. Chiu, I. M. *et al.* A Neurodegeneration-Specific Gene-Expression Signature of Acutely Isolated  
807 Microglia from an Amyotrophic Lateral Sclerosis Mouse Model. *Cell Rep.* **4**, 385–401 (2013).
- 808 21. Sun, S. *et al.* Translational profiling identifies a cascade of damage initiated in motor neurons  
809 and spreading to glia in mutant SOD1-mediated ALS. *Proc. Natl. Acad. Sci. U. S. A.* **112**, E6993-  
810 7002 (2015).
- 811 22. Vanlandewijck, M. *et al.* A molecular atlas of cell types and zonation in the brain vasculature.  
812 *Nature* **554**, 475–480 (2018).
- 813 23. Ma, Q., Ineichen, B. V., Detmar, M. & Proulx, S. T. Outflow of cerebrospinal fluid is  
814 predominantly through lymphatic vessels and is reduced in aged mice. *Nat. Commun.* **8**, (2017).
- 815 24. Kapoor, K. G., Katz, S. E., Grzybowski, D. M. & Lubow, M. Cerebrospinal fluid outflow: An  
816 evolving perspective. *Brain Res. Bull.* **77**, 327–334 (2008).
- 817 25. Tam, O. H. *et al.* Postmortem Cortex Samples Identify Distinct Molecular Subtypes of ALS:  
818 Retrotransposon Activation, Oxidative Stress, and Activated Glia. *Cell Rep.* **29**, 1164–1177.e5  
819 (2019).
- 820 26. Westeneng, H.-J. *et al.* Prognosis for patients with amyotrophic lateral sclerosis: development  
821 and validation of a personalised prediction model. *Lancet Neurol.* **17**, 423–433 (2018).
- 822 27. Braak, H. *et al.* Amyotrophic lateral sclerosis—a model of corticofugal axonal spread. *Nat. Publ.*  
823 *Gr.* **9**, 708–714 (2013).
- 824 28. Chiu, A. Y. *et al.* Age-Dependent Penetrance of Disease in a Transgenic Mouse Model of Familial  
825 Amyotrophic Lateral Sclerosis. *Mol. Cell. Neurosci.* **6**, 349–362 (1995).
- 826 29. Gille, B. *et al.* Inflammatory markers in cerebrospinal fluid : independent prognostic biomarkers  
827 in amyotrophic lateral sclerosis ? 1–9 (2019). doi:10.1136/jnnp-2018-319586
- 828 30. Turner, M. R., Goldacre, R., Talbot, K. & Goldacre, M. J. Cerebrovascular injury as a risk factor  
829 for amyotrophic lateral sclerosis: Table 1. *J. Neurol. Neurosurg. Psychiatry* **87**, 244–246 (2016).
- 830 31. Garton, F. C., Trabjerg, B. B., Wray, N. R. & Agerbo, E. Cardiovascular disease, psychiatric  
831 diagnosis and sex differences in the multistep hypothesis of amyotrophic lateral sclerosis. *Eur. J.*  
832 *Neurol.* 1–10 (2020). doi:10.1111/ene.14554
- 833 32. Rule, R. R., Schuff, N., Miller, R. G. & Weiner, M. W. Gray matter perfusion correlates with  
834 disease severity in ALS. *Neurology* **74**, 821–827 (2010).
- 835 33. Murphy, M. J. *et al.* Widespread cerebral haemodynamics disturbances occur early in  
836 amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler.* **13**, 202–9 (2012).
- 837 34. Van Laere, K. *et al.* Value of 18fluorodeoxyglucose-positron-emission tomography in



- 838 amyotrophic lateral sclerosis: a prospective study. *JAMA Neurol.* **71**, 553–61 (2014).
- 839 35. Wolak, T. Osteopontin - a multi-modal marker and mediator in atherosclerotic vascular disease.  
840 *Atherosclerosis* **236**, 327–37 (2014).
- 841 36. Chiocchetti, A. *et al.* Osteopontin Bridging Innate and Adaptive Immunity in Autoimmune  
842 Diseases. *J. Immunol. Res.* **2016**, 1–15 (2016).
- 843 37. Zhao, H. *et al.* The role of osteopontin in the progression of solid organ tumour. *Cell Death Dis.*  
844 **9**, 356 (2018).
- 845 38. Ding, J. *et al.* Large Perivascular Spaces Visible on Magnetic Resonance Imaging, Cerebral Small  
846 Vessel Disease Progression, and Risk of Dementia. *JAMA Neurol.* **74**, 1105 (2017).
- 847 39. Smeijer, D., Ikram, M. K. & Hilal, S. Enlarged Perivascular Spaces and Dementia: A Systematic  
848 Review. *J. Alzheimer's Dis.* **72**, 247–256 (2019).
- 849 40. Brown, R. *et al.* Understanding the role of the perivascular space in cerebral small vessel  
850 disease. *Cardiovasc. Res.* **114**, 1462–1473 (2018).
- 851 41. Skene, N. G. & Grant, S. G. N. Identification of vulnerable cell types in major brain disorders  
852 using single cell transcriptomes and expression weighted cell type enrichment. *Front. Neurosci.*  
853 **10**, 1–11 (2016).
- 854 42. Lerman, B. J. *et al.* Deletion of galectin-3 exacerbates microglial activation and accelerates  
855 disease progression and demise in a SOD1<sup>G93A</sup> mouse model of amyotrophic lateral sclerosis.  
856 *Brain Behav.* **2**, 563–575 (2012).
- 857 43. Gautier, L., Cope, L., Bolstad, B. M. & Irizarry, R. A. affy--analysis of Affymetrix GeneChip data at  
858 the probe level. *Bioinformatics* **20**, 307–315 (2004).
- 859 44. Durinck, S., Spellman, P. T., Birney, E. & Huber, W. Mapping identifiers for the integration of  
860 genomic datasets with the R/Bioconductor package biomaRt. *Nat. Protoc.* **4**, 1184–1191 (2009).
- 861 45. Ritchie, M. E. *et al.* limma powers differential expression analyses for RNA-sequencing and  
862 microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).
- 863 46. Rabin, S. J. *et al.* Sporadic ALS has compartment-specific aberrant exon splicing and altered cell-  
864 matrix adhesion biology. *Hum. Mol. Genet.* **19**, 313–328 (2009).
- 865 47. Harbor, B. Working with ALS Mice. *Jax* 1–28 (2009).
- 866 48. Ludolph, A. *et al.* A revision of the El Escorial criteria - 2015. *Amyotroph. Lateral Scler. Front.*  
867 *Degener.* **16**, 291–292 (2015).
- 868 49. van Vliet, E. A. *et al.* Expression and Cellular Distribution of P-Glycoprotein and Breast Cancer  
869 Resistance Protein in Amyotrophic Lateral Sclerosis Patients. *J. Neuropathol. Exp. Neurol.* **79**,  
870 266–276 (2020).
- 871 50. Casula, M. *et al.* Toll-like receptor signaling in amyotrophic lateral sclerosis spinal cord tissue.  
872 *Neuroscience* **179**, 233–243 (2011).
- 873 51. Huisman, M. H. B. *et al.* Population based epidemiology of amyotrophic lateral sclerosis using  
874 capture–recapture methodology. *J. Neurol. Neurosurg. & Psychiatry* **82**, 1165 LP-1170

- 875 (2011).
- 876 52. Brooks, B. R., Miller, R. G., Swash, M. & Munsat, T. L. El Escorial revisited: Revised criteria for  
877 the diagnosis of amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Other Mot. Neuron*  
878 *Disord.* **1**, 293–299 (2000).
- 879 53. Wallin, N. *et al.* The Swedish motor neuron disease quality registry. (2018).  
880 doi:10.1080/21678421.2018.1497065
- 881 54. Schwenk, J. M., Gry, M., Rimini, R., Uhlén, M. & Nilsson, P. Antibody Suspension Bead Arrays  
882 within Serum Proteomics. *J. Proteome Res.* **7**, 3168–3179 (2008).
- 883 55. Häggmark, A. *et al.* Plasma profiling reveals three proteins associated to amyotrophic lateral  
884 sclerosis. *Ann. Clin. Transl. Neurol.* **1**, 544–53 (2014).
- 885 56. Poesen, K. *et al.* Neurofilament markers for ALS correlate with extent of upper and lower motor  
886 neuron disease. *Neurology* **88**, 2302–2309 (2017).
- 887
- 888