Integrative gene–tissue microarray-based approach for identification of human disease biomarkers: application to amyotrophic lateral sclerosis

Lili C. Kudo1,3,† Liubov Parfenova1,4,†, Nancy Vl3,4, Kimbley Lau4, Justine Pomakian2, Paul Valdmanis5, Guy A. Rouleau6, Harry V. Vinters2, Martina Wiedau-Pazos1,‡ and Stanislav L. Karsten1,4,5,*‡

1Department of Neurology and 2Department of Pathology and Laboratory Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA 90095, USA, 3NeuroInDx, Inc., 1655 East 28th Street, Signal Hill, CA 90755, USA, 4Division of Neuroscience, Department of Neurology, Los Angeles Biomedical Research Institute and 5Department of Obstetrics and Gynecology, Harbor-UCLA Medical Center, Torrance, CA 90502, USA and 6Center for the Study of Brain Diseases, Notre-Dame Hospital, Centre de Recherche du CHUM, University of Montreal, Montreal, QC, Canada

Received May 7, 2010; Revised and Accepted June 3, 2010

Advances in genomics and proteomics permit rapid identification of disease-relevant genes and proteins. Challenges include biological differences between animal models and human diseases, high discordance between DNA and protein expression data and a lack of experimental models to study human complex diseases. To overcome some of these limitations, we developed an integrative approach using animal models, postmortem human material and a combination of high-throughput microarray methods to identify novel molecular markers of amyotrophic lateral sclerosis (ALS). We used laser capture microdissection coupled with microarrays to identify early transcriptome changes occurring in spinal cord motor neurons or surrounding glial cells. Two models of familial motor neuron disease, SOD1G93A and TAUP301L, transgenic mice were used at the presymptomatic stage. Identified gene expression changes were predominantly model-specific. However, several genes were regulated in both models. The relevance of identified genes as clinical biomarkers was tested in the peripheral blood transcriptome of presymptomatic SOD1 G93A animals using custom-designed ALS microarray. To confirm the relevance of identified genes in human sporadic ALS (SALS), selected corresponding protein products were examined by high-throughput immunoassays using tissue microarrays constructed from human postmortem spinal cord tissues. Genes that were identified by these experiments and located within a linkage region associated with familial ALS/frontotemporal dementia were sequenced in several families. This large-scale gene and protein expression study pointing to distinct molecular mechanisms of TAU- and SOD1-induced motor neuron degeneration identified several new SALS-relevant proteins (CNGA3, CRB1, OTUB2, MMP14, SLK, DDX58, RSPO2) and putative blood biomarkers, including Nefh, Prph and Mgll.

INTRODUCTION

Microarray-based approaches have revealed the regulatory complexity of neurological diseases and identified many useful disease modifiers (1–6). However, the translation of these results into new therapies has been hampered by a lack of suitable animal models, narrow experimental scope and limited
confirmation in human tissues. A high degree of technical and biological variability in gene expression microarray data (7) calls for a more comprehensive approach that will control for inherent differences between (i) animal models and human biology, (ii) RNA and protein expression, (iii) disease stages and (iv) cell-specific pathology. The development of such an inclusive approach is critical in translational research and has even greater significance for complex neurological diseases.

Amyotrophic lateral sclerosis (ALS) is a genetically heterogeneous neurodegenerative disease affecting motor neurons. It is the most common adult motor neuron disease (8). Mortality from this incurable disorder is often due to respiratory failure following an average of 3 years of progressive muscle weakness (9). Prominent clinical and pathological features include progressive dysfunction and loss of motor neurons in the anterior horns of the spinal cord, lower brainstem and cortex. Motor neurons of the upper brainstem and sacral area of the spinal cord are commonly spared (8). Many potential disease mechanisms have been implicated, including axonal dysfunction (10), protein aggregation (11–13), excitotoxicity (14,15), oxidative stress (16,17), mitochondrial dysfunction (18) and apoptotic mechanisms (19). However, the specific roles of these mechanisms in initiating and propagating neurodegeneration are largely unknown.

Familial ALS (FALS) represents ~10% of ALS cases and is most frequently inherited as an autosomal-dominant trait (8). In ~20% of FALS patients, a mutation in superoxide dismutase-1 (SOD1) can be found (20). The ubiquitous SOD1 protein converts superoxide radical anions to oxygen and hydrogen peroxide. More than 100 known SOD1 mutations (21) lead to a gain-of-function effect, which may contribute to an increased oxidative stress (16,22), altered copper metabolism (23), protein aggregation (24), excitotoxicity (14,15) or altered axonal transport (25). Patients with FALS caused by mutations in SOD1 can show comorbidity with frontotemporal dementia (FTD) and develop cognitive impairment, including apathy, inattention, verbal deficits and hypersexuality (26).

Several other genes have also been associated with FALS, including alsin, senataxin, VAPB, ANG, whereas TDP-43 has been linked to both familial and sporadic forms of ALS [for recent review see (27,28)]. Approximately 90% of ALS cases have no clear genetic cause, and the condition is referred to as sporadic amyotrophic lateral sclerosis [SALS (27)]. This genetic complexity is reflected in a wide range of pathophysiological features, including the not uncommon association with FTD (26,29).

At the cellular level, pathological signs of ALS include tau immunoreactive astrocytic and neuronal inclusions, suggesting that cognitive dysfunction in ALS may also reflect abnormal protein metabolism of the microtubule-associated protein (MAP), tau (30,31). This hypothesis is supported by the increased expression of Gsk3beta in some ALS patients, a major tau kinase that mediates tau hyperphosphorylation and aggregation (32). More evidence for direct tau involvement in ALS neuropathology is the association between ALS and FTDP-17 (FTD with Parkinsonism linked to chromosome 17), which is caused by TAU mutations (33–35). Thus, clinical, neuropathological, biochemical and genetic characteristics shared between FTD and ALS may suggest a common pathogenesis for motor neuron degeneration.

Large-scale microarray-based gene expression studies are well suited to clarify the role of specific disease mechanisms and to identify new cellular factors that contribute to motor neuron death in ALS. Previously published gene expression microarray studies in ALS animal models (36–42), cell cultures (43) and postmortem ALS central nervous system (CNS) tissues (44–49) have identified many genes potentially involved in the disease pathogenesis. However, few studies provided follow-up confirmation of the identified genes in human tissues or at the protein or cell level, in order to verify their functional significance. Such a limited validation partly explains the low overlap among different published genomic studies (<5% of all identified genes; Supplementary Material, Table S1) and emphasizes the need for careful cell-specific analysis and rigorous ‘post-array’ evaluation to identify the molecular factors relevant to the disease.

To identify cell-specific expression changes, we used laser capture microdissection (LCM) to isolate anterior horn motor neurons and surrounding cells. We then performed whole transcriptome analysis and used Chilibot and literature search to compare gene expression profiles in SOD1 and TAU mouse models, to determine common pathways for the development of ALS. Because ALS is a neurodegenerative disease that specifically affects motor neurons, these cells were the main investigative target. Global transcriptomes of glial cells surrounding the motor neurons were also assessed, since these cells have recently been implicated as triggers of neurodegeneration (50–53). Mouse experiments were performed at the pre-symptomatic stage, prior to the onset of cell loss, in order to reduce false-positive signals due to tissue-reactive changes. Identified gene expression changes were confirmed using semi-quantitative reverse transcriptase–polymerase chain reaction (sqRT–PCR) and immunohistochemistry (IHC).

Next, tissue microarrays (TMAs) in ALS postmortem tissues were used to identify proteins that are products of identified candidate genes, which may be linked to specific motor neuron vulnerability (Fig. 1). Selected corresponding protein candidates were subjected to high-throughput immunohistochemical analysis using TMAs constructed from human postmortem SALS and control spinal cord specimens. Genes that were identified as valid candidates for motor neuron vulnerability by these experiments and located within a linkage region associated with FALS with FTD (ALS/FTD) were sequenced in several families. Finally, to test whether any of the identified SOD1 motor neuron or glia-specific genes may be detected in the peripheral blood and therefore be used as clinical biomarkers, ALS-specific microarrays were used for the analysis of gene expression in the peripheral blood of presymptomatic SOD1G93A mice.

This study establishes several crucial connections between: (i) mouse and human motor neuron pathologies; (ii) familial and sporadic ALS; (iii) early and terminal stages of the disease; (iv) RNA and protein expression and (v) relevance of the identified motor neuron and glia-specific gene expression changes as peripheral blood biomarkers. This is the first report of large-scale gene and protein expression analyses of spinal cord motor neuron and glial cell populations that identify overlapping gene and protein expression changes in mouse models of familial motor neuron disease and human SALS.
Identification of SOD1<sup>G93A</sup> and TAUP<sup>P301L</sup> presymptomatic gene expression changes

We used laser capture microdissection and high-density oligonucleotide microarrays to identify genes for which expression is altered in motor neuron bodies and in the surrounding glia of presymptomatic SOD1<sup>G93A</sup> and TAU<sup>P301L</sup> transgenic animals (Fig. 1). These techniques identified 251 transcripts representing 186 known genes for which expression was altered in at least one of the four comparisons (Supplementary Material, Table S2). Gene ontology (GO) analysis using DAVID revealed that the biological processes most enriched in motor neurons and in surrounding cells are those associated with protein modification/phosphorylation, signaling, muscle contraction regulation, stress responses, the immune system and cell communication (Figs 2 and 3).
Twelve transcripts including eight known genes were regulated in both SOD1G93A and TAU P301L transgenic animals (Table 1). Five (Cnga3, Crb1, Dnahc7b, Otub2 and Slk) were detected in motor neurons and three (Dusp7, Mrm1 and Zcchc17) in surrounding cells (Table 1).

In total, 79 and 53 known genes were identified in SOD1G93A motor neurons and surrounding cells, respectively (Supplementary Material, Table S2). Ten of these genes have been previously implicated in ALS or motor neuron disease or motor neuron metabolism (Fig. 2). Twenty-two genes (17%) had been identified in other recent microarray-based ALS studies (Supplementary Material, Table S1).

Thirty-nine and 24 known genes were regulated in motor neurons and surrounding cells of TAU P301L mice, respectively (Supplementary Material, Table S2). Among them 12 genes (22%) were found in previous ALS microarray studies (Supplementary Material, Tables S1 and S2). GO analysis revealed that genes involved in the regulation of muscle contraction including calcitonin (Calca, TAUmn↑) and guanylate cyclase 1, soluble, alpha 3 (Gucy1a3, TAUmn↓) were significantly enriched in TAU G93A motor neurons (Fig. 3C), pointing to early deregulation of neuromuscular innervations in these mice. TAU P301L glial cells demonstrated moderate enrichment for regulated genes related to apoptosis, immune system response and post-translational protein modification (Fig. 3).

**Confirmation of cell-specific microarray gene expression data**

Semi-quantitative RT–PCR experiments of a few selected motor neuron-specific genes confirmed the microarray findings. All tested motor neuron-specific genes (Cnga3, Otub2, Slk, Mmp14 and Ccl19) demonstrated the same pattern of regulation as observed in microarrays, when LCM-collected motor neurons from both SOD1G93A and TAU P301L were used for RT–PCR (Fig. 3).

To investigate the cell specificity and relevance of the microarray-identified transcriptional changes at the protein level in mouse tissues, immunohistochemical experiments on additional SOD1G93A and TAU P301L mouse spinal cord tissues were performed for three selected proteins, Cnga3 (SOD1mn↑; TAUmn↑), Ddx58 (SOD1mn↑) and Mmp14 (TAUmn↓) (Fig. 3). The selection of proteins was done based on two criteria: ‘low-end’ microarray ratios corresponding to a 1.5-fold change and the availability of the commercial antibodies (Table 1; Supplementary Material, Table S3). The low end of microarray detection was chosen to test the sensitivity of the selected cut-off. A noticeable increase in Cnga3 protein expression in anterior horn motor neurons in both SOD1G93A and TAU P301L transgenic mice, increases in Ddx58 protein expression in motor neurons of SOD1G93A mice and a decrease in Mmp14 protein in motor neurons of TAU P301L mice were observed, confirming the LCM-based microarray data (Fig. 4; Table 1). These results provide additional assurance that even low-end microarray-identified gene expression changes may directly be reflected in protein levels.

**Analysis of selected candidates in human SALS postmortem spinal cord tissues identifies several new ALS-relevant proteins**

One goal of this study was translation of gene expression findings in multiple transgenic mouse models of familial motor neuron disease to human SALS. We hypothesized that gene expression changes identified in more than one genetic FALS model may represent global markers of overall motor neuron vulnerability rather than a single-gene effect. To test whether any of the regulated genes identified are specific for human SALS, we used TMAs constructed of several clinically relevant CNS regions from SALS and normal control subjects (Supplementary Material, Fig. S1, Table S4). Immunohistochemical analyses were performed against the protein products of 9 selected gene candidates, namely CNGA3, CRB1, OTUB2, MMP14, CCL19, S100A9, SLK, DDX58 (Fig. 5, Table 2) and DNAHC7b, for which antibodies against the corresponding human protein homolog were utilized (Supplementary Material, Fig. S2). Priority was given to genes regulated in both mouse models and for which antibodies against corresponding proteins are commercially available.

Consistent with our mouse microarray and protein data, we found that CRB1 (SOD1mn↑, TAUmn↑), CNGA3 (SOD1mn↑, TAUmn↑) and OTUB2 (SOD1mn↑, TAUmn↑) protein levels were significantly increased in SALS anterior horn motor neurons, and the level of MMP14 (TAUmn↓) protein was significantly decreased (Fig. 5). Rsps2 oppositely regulated in both SOD1G93A and TAU P301L motor neurons (SOD1mn↓; TAUmn↑) showed decreased levels of corresponding human protein homolog in SALS tissues in agreement with SOD1G93A microarray data (Fig. 6). We observed contrasting results of mouse and human studies with respect to DDX58 (SOD1mn↑) and SLK (SOD1mn↑, TAUmn↑), which were upregulated in both presymptomatic mouse tissues but significant downregu-
lated in the human SALS tissue samples (Fig. 5). Similarly, transcriptional activation of dual specificity phosphatase 7 (Dusp7; SOD1\(^{G93A}\); TAU\(^{G93A}\)) and tissue inhibitor of metalloproteinases 2 (Timp2; SOD1\(^{G93A}\); TAU\(^{G93A}\)) in glia of presymptomatic mice was not in agreement with human postmortem data that demonstrated

---

**Table 1.** Gene expression changes identified in both SOD1\(^{G93A}\) and TAU\(^{P301L}\) motor neurons and surrounding glial cells.

<table>
<thead>
<tr>
<th>Gene ID</th>
<th>SODmn</th>
<th>SODgl</th>
<th>TAUmn</th>
<th>TAUgl</th>
<th>Symbol</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>12790</td>
<td>1.5</td>
<td>1.0</td>
<td>1.6</td>
<td>−1.3</td>
<td>Cnga3</td>
<td>Cyclic nucleotide-gated channel alpha 3</td>
</tr>
<tr>
<td>170788</td>
<td>2.3</td>
<td>1.1</td>
<td>4.5</td>
<td>1.1</td>
<td>Cri1</td>
<td>Crlms homolog 1 (Drosophila)</td>
</tr>
<tr>
<td>227058</td>
<td>2.3</td>
<td>2.3</td>
<td>2.4</td>
<td>1.3</td>
<td>Dnahc7b</td>
<td>Dynein, axonemal, heavy-chain 7B</td>
</tr>
<tr>
<td>68149</td>
<td>1.5</td>
<td>1.5</td>
<td>1.6</td>
<td>1.6</td>
<td>Otub2</td>
<td>OTU domain, ubiquitin aldehyde binding 2</td>
</tr>
<tr>
<td>20874</td>
<td>1.9</td>
<td>1.1</td>
<td>4.9</td>
<td>1.2</td>
<td>Slk</td>
<td>STE20-like kinase (yeast)</td>
</tr>
<tr>
<td>239405</td>
<td>−1.7</td>
<td>1.2</td>
<td>1.9</td>
<td>1.9</td>
<td>Rspo2</td>
<td>R-spondin 2 homolog (Xenopus laevis)</td>
</tr>
<tr>
<td>235584</td>
<td>1.3</td>
<td>2.0</td>
<td>1.3</td>
<td>5.2</td>
<td>Dusp7</td>
<td>Dual specificity phosphatase 7</td>
</tr>
<tr>
<td>217038</td>
<td>1.0</td>
<td>1.6</td>
<td>1.0</td>
<td>2.6</td>
<td>Mrm1</td>
<td>Mitochondrial rRNA methyltransferase 1</td>
</tr>
<tr>
<td>619605</td>
<td>1.1</td>
<td>1.9</td>
<td>−1.1</td>
<td>2.7</td>
<td>Zcch17</td>
<td>Zinc finger, CCHC domain containing 17</td>
</tr>
</tbody>
</table>

Average fold changes between transgenic mice and nontransgenic controls are shown. Official Entrez Gene symbol and numbers are used (184). For the full list of genes identified in this study, see Supplementary Material. Boldface and italic values indicate significantly upregulated (bold) or downregulated (italic) genes. SOD1mn, genes altered in SOD1\(^{G93A}\) motor neurons; SOD1gl, genes altered in the cells surrounding SOD1\(^{G93A}\) motor neurons; TAUmn, genes altered in TAU\(^{P301L}\) motor neurons; TAUgl, genes altered in the cells surrounding TAU\(^{P301L}\) motor neurons.

---

**Figure 3.** Gene ontology analysis identifies significant enrichment (EASE score <0.05) of genes involved in the following biological processes: protein modification in both mouse models (A); signaling in SOD1\(^{G93A}\) motor neurons (B); muscle contraction in TAU\(^{P301L}\) motor neurons (C) and enrichment for genes involved in stress response (D); immune response (E) and cell communication in SOD1\(^{G93A}\) glial cells (F). All genes, all gene expression changes identified in at least one comparison of the current microarray study; All mn, motor neuron-specific gene expression changes; SOD mn and TAU mn, SOD\(^{G93A}\)- and TAU\(^{P301L}\)-specific motor neuron expression changes, respectively; All glia, glia-specific gene expression changes; SOD glia and TAU glia, SOD\(^{G93A}\)- and TAU\(^{P301L}\)-specific glia expression changes, respectively. The numbers of genes and corresponding percent of each biological process relative to a total number of genes in a specific comparison are shown next to the bars on the right.
a sharp decrease for all three proteins in SALS spinal cord (Fig. 6). No changes or contradictory results were detected for CCL19 (TAUmn↓, TAUgl↓) and S100A9 (TAUmn↓) proteins in SALS samples. Interestingly, the two types of TMA analysis for CCL19 demonstrated opposite results. CCL19 protein expression based on the number of cells expressing CCL19 (TMA2) agreed with murine microarray data, whereas TMA analysis based on the calculation of average motor neuron-specific signal (TMA1) did not confirm the mouse data (Fig. 5).
Figure 5. Immunohistochemical analyses of human SALS tissue microarrays (SALS–TMA) confirm altered protein expression of several genes identified in presymptomatic SOD1<sup>G93A</sup> and TAU<sup>P301L</sup> transgenic mice, detected by the LCM-based microarray approach. (A) Representative tissue cores from human post-mortem anterior horns demonstrating protein expression changes in SALS and correlation with mouse microarray data. The green panel indicates protein expression changes that are in agreement with mouse array data. The yellow panel, representing CCL19 and S100A9, indicates protein changes that either do not parallel the LCM data or only partially support them (in the case of CCL19). The red panel, representing SLK and DDX58, demonstrates protein expression changes in SALS tissues opposite to the changes observed for mouse microarray data. Black arrows summarize the direction of protein expression in SALS motor neurons, where ‘↑’ indicates elevated and ‘↓’ decrease expression in SALS motor neurons. Bar graphs on the left (B) show the average folds expression changes between SALS and control (TMA 1 and 2) or microarray result expressed as average fold change between transgenic mouse and control (MA TAU and SOD). An asterisk indicates $P \leq 0.05$ (two-tailed Student’s $t$-test). ‘TMA1’ and ‘TMA2’ show the results of two different TMA data analysis (see Materials and Methods). MA, microarray result; TAU, microarray results from TAU<sup>P301L</sup> motor neurons; SOD, microarray results from SOD1<sup>G93A</sup> motor neurons. Scale bars correspond to 200 μm.
Expression changes in the peripheral blood of presymptomatic SOD1G93A mice
To test whether any of the identified genes are expressed in the peripheral blood of presymptomatic SOD1G93A mice, we have constructed a targeted custom microarray (ALS-array) harboring genes identified in the initial genomic screen and genes demonstrating a trend for regulation (Supplementary Material, Table S2). Analysis of total RNA isolated from the peripheral blood of presymptomatic SOD1G93A mice using ‘ALS-array’ revealed that the expression of 13 out of 132 (10%) SOD1G93A motor neuron and glia-specific genes is significantly altered (Table 3). Three of these genes, namely monoglyceride lipase (Mgll; SODgl↑, SODb↑), neurofilament, heavy polypeptide (Nefh; SODgl↓, SODb↓) and Peripherin (Prph; SODgl↓, SODB↓), appear to play a direct functional role in ALS pathogenesis (Table 3).

Chilibot and literature searching for interactions between SOD1 and TAU
Our analyses found few similarities in gene expression among motor neurons and glia of SOD1G93A and TAU P301L mice, indicating that their phenotypic similarities belie their differences. Chilibot analysis and literature searching were undertaken to determine whether interactions between TAU and SOD1 could explain the similar phenotype. References to TAU were scanned for the genes identified in the SOD1G93A mouse model, and vice versa (SOD1 references for genes found with the TAU P301L model). The procedure identified six genes altered in the mutant SOD1G93A mice that have known roles in TAU biology: Klk6, Rasd2, Rps6kb1, Prph, Pmp22 and Nefh (Fig. 2). Similarly, we found four genes regulated in TAU P301L mice that had been implicated in SOD1 metabolism (Casp9, Calca, Ndr1 and Htr3a; Fig. 2), suggesting a possible functional link between SOD1G93A and TAU P301L induced motor neuron disease.

DISCUSSION
We investigated whether TAU- and SOD1-induced vulnerability to motor neuron degeneration involves similar regulatory pathways. In order to elucidate these pathways, and to identify putative functional biomarkers of SALS, we employed a sequential unbiased approach that combined gene and TMAs with analyses of animal and human tissues at both RNA and protein levels. To our knowledge, this is the first study of a neurodegenerative disease that identifies and confirms a large number of gene candidates at the protein level in both mouse
models and human tissue. The experimental design offers several compelling advantages over typical microarray-based studies. We employed LCM for cell-specific analyses of the motor neuron and glial transcriptomes, in order to tackle the identified morphological phenotype. We used two genetically distinct mouse models and human tissue. The experimental design offers several compelling advantages over typical microarray-based studies. We employed LCM for cell-specific analyses of the motor neuron and glial transcriptomes, in order to tackle the mechanisms of anterior horn motor neuron degeneration in ALS. Parallel analysis of the immediate cellular environment surrounding motor neurons was also chosen, because of the established importance of glial cells in neurodegeneration.

To minimize the identification of false positives related to progressive neurodegeneration in advanced stages of the disease, and to increase our chances for identifying ALS biomarkers, collection of cellular material for microarray analyses was done prior to the appearance of any detectable morphological phenotype. We used two genetically distinct but phenotypically similar models, SOD1<sup>G93A</sup> and TAUP<sup>301L</sup> transgenic mice, to further increase the likelihood of identifying genes relevant to ALS that are not limited to a single genetic cause (59,60). Microarray-identified transcriptome changes were independently validated at RNA and protein levels, selected genes were evaluated in human SALS using TMA technology (Table 2) and the relevance of identified changes as clinical biomarkers was tested in the peripheral blood transcriptome of presymptomatic SOD1<sup>G93A</sup> mice using custom-designed microarrays (Table 3).

Notably, several cytoskeletal proteins previously implicated in motor neuron disease were identified in SOD1<sup>G93A</sup> and/or TAUP<sup>301L</sup> mice. SOD1<sup>G93A</sup>-specific genes included peripherin (Prph; SOD1mn<uparrow>), heavy neurofilament protein (Nefh; SOD1gl<downarrow>), cathepsin H (Ctsb; SOD1gl↑) and adenosine deaminase, RNA-specific, B1 (Adarb1; SOD1gl↑). Peripherin encodes a type III intermediate filament protein that is

### Table 2. Summary of ‘post-array’ studies

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Name</th>
<th>SODmn</th>
<th>SODgl</th>
<th>TAUmn</th>
<th>TAUgl</th>
<th>Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cnga3</td>
<td>Cyclic nucleotide-gated channel alpha 3</td>
<td>1.5</td>
<td>1.0</td>
<td>1.6</td>
<td>−1.3</td>
<td>RT, IHC, TMA</td>
</tr>
<tr>
<td>Crb1</td>
<td>Crumbs homolog 1 (Drosophila)</td>
<td>2.3</td>
<td>1.4</td>
<td>4.5</td>
<td>1.1</td>
<td>TMA</td>
</tr>
<tr>
<td>Dnabh7b</td>
<td>Dynacin, axonemal, heavy-chain 7B</td>
<td>2.3</td>
<td>2.3</td>
<td>2.4</td>
<td>1.3</td>
<td>TMA</td>
</tr>
<tr>
<td>Otub2</td>
<td>OTU domain, ubiquitin aldehyde binding 2</td>
<td>1.5</td>
<td>1.5</td>
<td>1.5</td>
<td>1.6</td>
<td>RT, TMA</td>
</tr>
<tr>
<td>Rspo2</td>
<td>R-spondin 2 homolog (Xenopus laevis)</td>
<td>−1.7</td>
<td>1.2</td>
<td>−1.9</td>
<td>1.9</td>
<td>WB</td>
</tr>
<tr>
<td>Slk</td>
<td>STE20-like kinase (yeast)</td>
<td>1.9</td>
<td>1.1</td>
<td>4.9</td>
<td>1.2</td>
<td>RT, TMA</td>
</tr>
<tr>
<td>Dusp7</td>
<td>Dual specificity phosphatase 7</td>
<td>1.3</td>
<td>2.0</td>
<td>1.3</td>
<td>5.2</td>
<td>WB</td>
</tr>
<tr>
<td>Ddx58</td>
<td>DEAD (Asp-Glu-Ala-Asp) box polypeptide 58</td>
<td>1.5</td>
<td>1.0</td>
<td>1.7</td>
<td>1.1</td>
<td>IHC, TMA</td>
</tr>
<tr>
<td>Timp2</td>
<td>Tissue inhibitor of metalloproteinase 2</td>
<td>1.0</td>
<td>1.6</td>
<td>1.1</td>
<td>1.2</td>
<td>WB</td>
</tr>
<tr>
<td>Nefh</td>
<td>Neurofilament, heavy polypeptide</td>
<td>−1.2</td>
<td>−1.7</td>
<td>1.2</td>
<td>1.5</td>
<td>TMA</td>
</tr>
<tr>
<td>Cc19,n</td>
<td>Chemokine (C-C motif) ligand 19</td>
<td>1.4</td>
<td>2.0</td>
<td>−3.7</td>
<td>−3.2</td>
<td>RT, TMA, Seq</td>
</tr>
<tr>
<td>Dctn3</td>
<td>Dynactin 3</td>
<td>1.0</td>
<td>−1.1</td>
<td>−1.7</td>
<td>1.0</td>
<td>Seq</td>
</tr>
<tr>
<td>Mmp14</td>
<td>Matrix metallopeptidase 14</td>
<td>1.7</td>
<td>1.7</td>
<td>−1.6</td>
<td>−1.1</td>
<td>RT, IHC, TMA</td>
</tr>
<tr>
<td>S100A9</td>
<td>S100 calcium binding protein A9 (calgranulin B)</td>
<td>3.4</td>
<td>1.4</td>
<td>−2.2</td>
<td>1.3</td>
<td>TMA</td>
</tr>
<tr>
<td>Timp3</td>
<td>Tissue inhibitor of metalloproteinase 3</td>
<td>1.4</td>
<td>1.4</td>
<td>1.7</td>
<td>1.6</td>
<td>WB</td>
</tr>
</tbody>
</table>

Selected genes (n = 15) were used for RT-PCR (RT) and immunohistochemistry (IHC) in mouse tissues, tissue microarray (TMA) and western blot (WB) analysis in SALS postmortem samples, and sequencing in ALS-FTD families (Seq). Boldface and italic values indicate significantly upregulated (bold) or downregulated (italic) genes.

### Table 3. Gene expression changes identified in SOD1<sup>G93A</sup> peripheral blood

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Name</th>
<th>SODmn</th>
<th>SODgl</th>
<th>TAUmn</th>
<th>TAUgl</th>
<th>SODb</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ccpg1</td>
<td>Cell cycle progression 1</td>
<td>−2.0</td>
<td>−1.1</td>
<td>1.1</td>
<td>1.1</td>
<td>−2.0</td>
<td>0.008</td>
</tr>
<tr>
<td>Cby56l</td>
<td>Cytochrome b-561</td>
<td>−1.6</td>
<td>−1.2</td>
<td>1.1</td>
<td>1.7</td>
<td>−2.7</td>
<td>0.005</td>
</tr>
<tr>
<td>Dclk1</td>
<td>Doublecortin-like kinase 1</td>
<td>−1.7</td>
<td>1.3</td>
<td>1.2</td>
<td>2.6</td>
<td>−2.0</td>
<td>0.028</td>
</tr>
<tr>
<td>Efhh1</td>
<td>EF-hand domain (C-terminal) containing 1</td>
<td>−1.7</td>
<td>−1.3</td>
<td>−1.1</td>
<td>1.0</td>
<td>−2.2</td>
<td>0.005</td>
</tr>
<tr>
<td>Kctd8</td>
<td>Potassium channel tetramerisation domain containing 8</td>
<td>−1.7</td>
<td>1.1</td>
<td>1.5</td>
<td>1.7</td>
<td>−2.2</td>
<td>0.013</td>
</tr>
<tr>
<td>Mgll</td>
<td>Monoacyl glyceride lipase</td>
<td>1.0</td>
<td>2.1</td>
<td>−1.2</td>
<td>1.2</td>
<td>1.5</td>
<td>0.111</td>
</tr>
<tr>
<td>Nefh</td>
<td>Neurofilament, heavy polypeptide</td>
<td>−1.2</td>
<td>−1.7</td>
<td>1.2</td>
<td>1.5</td>
<td>−2.8</td>
<td>0.001</td>
</tr>
<tr>
<td>Prph</td>
<td>Peripherin</td>
<td>−1.7</td>
<td>−1.1</td>
<td>1.1</td>
<td>1.4</td>
<td>−3.0</td>
<td>0.002</td>
</tr>
<tr>
<td>Rspo2</td>
<td>R-spondin 2</td>
<td>−1.7</td>
<td>1.2</td>
<td>1.9</td>
<td>1.9</td>
<td>−2.8</td>
<td>0.008</td>
</tr>
<tr>
<td>Stk32a</td>
<td>Serine/threonine kinase 32A</td>
<td>−1.6</td>
<td>1.2</td>
<td>1.5</td>
<td>1.0</td>
<td>−2.0</td>
<td>0.007</td>
</tr>
<tr>
<td>Usp46</td>
<td>Ubiquitin specific peptidase 46</td>
<td>−1.2</td>
<td>−1.5</td>
<td>1.3</td>
<td>1.2</td>
<td>−2.0</td>
<td>0.002</td>
</tr>
<tr>
<td>Vezt</td>
<td>Vezatin</td>
<td>−1.6</td>
<td>1.0</td>
<td>1.2</td>
<td>1.0</td>
<td>−2.2</td>
<td>0.004</td>
</tr>
<tr>
<td>Vitrin</td>
<td>Vitrin</td>
<td>−1.9</td>
<td>−1.4</td>
<td>1.1</td>
<td>1.1</td>
<td>−2.8</td>
<td>0.028</td>
</tr>
</tbody>
</table>

Average fold changes between transgenic mice and nontransgenic controls are shown. Raw P-values are shown for the blood transcriptome analysis. Official Entrez Gene symbol and numbers are used (183). For details, see Supplementary Material and the MIAME report. Boldface and italic values indicate significantly upregulated (bold) or downregulated (italic) genes.
associated with motor neuron degeneration in ALS (61–63) and aggregates in the motor neurons of SOD1 transgenic mice (63). Mutations in the peripheral myelin protein 22 (Pmp22; SOD1mn↓) cause Charcot–Marie–Tooth disease Type IA, Dejerine–Sottas syndrome and hereditary neuropathy with liability to pressure palsies (64). Similar to Prph, Pmp22 is prone to aggregation (64). Mutations in heavy neurofilament protein (Nefh; SOD1gl↓) are found in some familial ALS patients (65,66), and Nefh transgenic mice develop motor neuron degeneration (67–70). Elevated levels of cathespin H (Ctsb; SOD1gl↑), a lysosomal cysteine protease, were recently identified in the astrocytes and microglia of the spinal cord ventral horn of presymptomatic SOD1G93A↓,TauP301L↓ transgenic mice (37), and are compatible with our microarray data (Supplementary Material, Tables S1 and S2). Ctsb activation is associated with the reaction against abnormally misfolded and aggregated proteins, including PRPH, PMP22, SOD1 and TAU (71–74). Adenosine deaminase, RNA-specific, B1 (Adarb1; SOD1gl↑), a regulator of AMPA receptor maturation and its Ca2+ permeability (75) directly affects selective motor neuron death in SALS (76,77).

Three genes, caspase 9 (Casp9; Taumn↓), nucleolin (Ncl; TauUm↓) and serpin peptidase inhibitor, clade I, member 1 (neuroserpin/serpinp1; Taugl↑), have known roles in Tau biology (78–80). Casp9 is activated prior to Tau accumulation and is involved in Tau protein cleavage (80). Ncl is an early marker of neurofibrillary tangle formation [NFTs (79)]. Mutations in neuroserpin, an important regulator of axonal growth, result in familial encephalopathy with dementia characterized by neuroserpin inclusion bodies (78).

Shared mechanisms of SOD1- and Tau-induced motor neuron degeneration

Simultaneous transcriptome analysis of SOD1- and Tau-induced motor neuron degeneration revealed only eight genes (Cnga3, Crb1, Dnahe7b, Otub2, Slk, Dusp7, Mrm1 and Zcchc17) similarly regulated in both models (Table 1). While regulation of Cnga3, Crb1, Otub2, Dusp7, Mrm1 and Zcchc17 has not been previously described in motor neurons, two genes, axonemal dynein heavy-chain 7b (Dnahe7b; Sod1mn↑; Tauum↓) and Ste-20 like kinase (Slk; Sod1mn↑; Tauum↑), are known to be involved in motor neuron function. Specifically, Dnahe7b is involved in antero- and retrograde transport in cellular processes related to motor neuron disease. Moreover, mutations in dynein—a protein associated with Dnahe7b—are found in some familial ALS cases (81,82), and interactions of dynein with mutated SOD1 in transgenic mice have been reported (83,84). Slk is a microtubule-associated kinase expressed at the neuromuscular junction that is linked to motor neuron degeneration through microtubule-based axonal transport (85–87). The small degree of overlap in expressed genes between the two mouse models suggests that the mechanism of familial SOD1- and Tau-induced neurodegeneration, in spite of some phenotypic similarities, is most likely different, with very few shared processes taking place at the early stage of disease. This may also imply that genes regulated in common may mark metabolic and regulatory processes involved in early motor neuron dysfunction in general, and therefore represent potential biomarker candidates for disease pathogenesis and progression in SALS as well as familial ALS.

To test whether regulation of identified factors occurs at the protein level in clinical human disease, we applied high-throughput analysis of postmortem SALS tissues using TMA technology (Supplementary Material, Fig. S1). TMAs were initially described by Kononen et al. (88) and since then have been applied in cancer research and in the study of neurological disorders (89). Our microarray experiment identified five genes that were upregulated in motor neurons of both SOD1G93A↓ and TauP301L↓ transgenic mice (Cngas3, Crb1, Dnahe7b, Slk and Otub2, Table 1). TMA analysis confirmed mouse microarray data for three genes, CNGA3, CRB1 and OTUB2, which were upregulated in SALS motor neurons (Fig. 5). SLK showed regulation in a direction opposite to that found with microarrays, and analysis of dynein complex (Dnahe7b) did not reveal any changes in SALS tissues (Supplementary Material, Fig. S2).

In addition to genes regulated in common, both models demonstrated evidence for potentially overlapping mechanisms of SOD1G93A↓- and TauP301L↓-induced motor neuron degeneration. Chilibot analysis and literature searching identified six genes related to Tau biology that were altered in mutant SOD1G93A↓ mice. Specifically, kallikrein-related peptidase 6, or neurosin (Klk6; Sod1mn↓), is a marker of Alzheimer’s disease and is involved in generation of amyloidogenic fragments from the amyloid precursor protein (90). Ribosomal protein S6 kinase (Rps6kb1; Sod1mn↓) phosphorylates Tau protein and enhances its aggregation (91,92). RASD family, member 2 (Rasd2; Sod1mn↑) is related to Rasd1 that inhibits expression of GSK3 beta, which in turn directly phosphorylates Tau protein (93). Peripherin (Prph; Sod1mn↓) and neurofilament, heavy polypeptide (Nefh; Sod1gl↓) along with Tau, are cytoskeletal proteins. Prph and Tau aggregates are both found in ALS patients (94). Nefh serves as a pathological chaperone in the formation of intraneuronal Tau aggregates (95). A mouse knockout for peripheral myelin protein 22 (Pmp22; Sod1mn↓) is characterized by deficient axonal transport, neurofilament phosphorylation, decreased microtubule stability and hyperphosphorylation of axonal MAPs, including Tau (96,97).

In TauP301L↓ mice, a search for regulated genes associated with SOD1 function revealed four genes previously implicated in SOD1 metabolism (Fig. 2). Caspase 9 (Casp9; TauUm↓) downregulates SOD1 through p53 activation (98,99). Low levels of Casp9 were previously detected in motor neurons in presymptomatic SOD1 mutant mice (41). A common response to elevated oxidative stress in SOD1 knockout mice is the activation of p53 target genes, which include N-Myc downstream regulated 1 (Ndr1; TauGl↓). Interestingly, mutations in NDRG1 are responsible for Charcot–Marie–Tooth neuropathy type 4D, characterized by severe demyelination and axonal loss due to impaired axon–glia interactions (Online Mendelian Inheritance in Man; http://www.ncbi.nlm.nih.gov/omim; #601455).

Intriguingly, one protein, R-spondin 2 was regulated in both models but in the opposite orientations (Rspo2; Sod1mn↓; TauUm↑). Rspo2 is an activator of the Wnt pathway through direct interaction with beta-catenin (100,101).
Protein analysis demonstrated a significant decrease of RSPO2 in SALS tissues in agreement with SOD1\textsuperscript{G93A} microarray data (Table 1; Fig. 6). In addition, significantly decreased levels of Rspo2 were detected in presymptomatic SOD1\textsuperscript{G93A} peripheral blood, suggesting that it may be a feasible clinical biomarker of early motor neuron degeneration (Table 3).

**Mouse–human ALS transcriptomes**

Fundamental differences between mouse and human physiology complicate the direct comparison of mouse models with human diseases, such as ALS. Several additional factors contribute to the difficulty of this task. First, the majority of human ALS cases are sporadic and most likely caused by a combination of multiple genetic and environmental factors. Existing ALS mouse models mimic familial ALS cases, which originate from mutations in a specific gene, such as SOD1. This ‘single-gene effect’ may be alleviated to a certain extent by the use of additional models with a similar phenotype, caused by a different genetic defect, such as TAU in the current study. This approach represents a powerful way to identify even a small number of functionally important, overlapping genes and their protein products that contribute to motor neuron vulnerability.

In 11 recent ALS mouse or human microarray studies, including the current work, 1140 known unique genes have been identified. These experiments revealed low overlap among the identified genes, a common phenomenon in microarray studies. Explanations include the use of different probes in different microarray platforms, different mouse models, variability in motor neuron collection techniques, age and sex of the animals and labeling protocols (102). Only 95 genes (8.3%) were found in more than one study, with 61 genes demonstrating the same direction of gene expression change (5.4%; Supplementary Material, Table S1).

Human ALS postmortem spinal cord tissues contain a reduced number of motor neurons, and the remaining motor neurons are expected to be in various stages of degeneration. Terminal stage of the disease is associated with multiple effects on global gene expression heavily contaminating the pool of potential disease modifiers identified in the microarray studies with genes misregulated due to reactive changes, such as cell death and reactive immune responses. These ‘false positives’ may be filtered out by comparison of human postmortem data with the mouse presymptomatic transcriptome. Comparisons should be done with careful consideration of methodologies, material and specific mouse models used. For example, changes identified in whole tissues may reflect the changes in cellular composition associated with neurodegeneration (1,7,103,104). Therefore, cell-specific analysis is required for identification of motor neuron-specific changes. Several attempts have been made to identify motor neuron-specific gene expression changes in mouse models (36,39,41) and human SALS (46) using LCM (Supplementary Material, Table S1). One study looked at the human SALS postmortem motor neuron transcriptome using BD Atlas Glass Microarrays Human 1.0 and 3.8 [Clontech; (46)]. Out of 185 known SOD1- or TAU-associated genes identified here, 38 genes (34%) were present on the Clontech arrays and 23 genes (60%) were found to be regulated in human SALS motor neurons—with 10 of these genes demonstrating a trend for regulation similar to that found in our study ([46], Supplementary Material, Table S2). Another attempt to correlate human SALS postmortem data with the SOD1\textsuperscript{G93A} mouse phenotype was recently performed (48). Although the initial screening experiment used low-density cDNA arrays (1176 probes), postmortem tissues and a small group of SALS subjects (n = 4), six genes identified in human SALS tissues were cross-confirmed in symptomatic (4 months) SOD1\textsuperscript{G93A} mice.

**Axonal transport dysfunction**

Defects in axonal transport contribute to the pathogenesis of many neurodegenerative disorders, including ALS and FTD (105,106). Dynein/dynactin is the centerpiece protein complex governing retrograde axonal transport in motor neurons (105,107,108). Its inhibition results in late onset motor neuron degeneration (109). Mutations in the largest subunit of dynactin, p150, cause rare cases of familial ALS (82), and mice with mutated dynein heavy chain also develop motor neuron degeneration (110). Moreover, ALS patients demonstrate deficiencies in mitochondria retrograde transport (111–113).

Several axonal transport-specific proteins identified here to be regulated in both SOD1\textsuperscript{G93A} and TAU\textsuperscript{P301L} motor neurons also showed a trend for upregulation in the surrounding cells (Table 1). These included dynein and axonemal, heavy-chain 7B protein (Dnahc7b, SODmn\textsuperscript{+}, TAUmn\textsuperscript{+}). However, TMA immunohistochemistry did not confirm upregulation of the dynein protein complex in SALS motor neurons (Supplementary Material, Fig. S4), suggesting that the identified changes may be gene or disease stage-specific. Indeed, dysfunction of dynein-dependent transport was reported to be specifically associated with mutations in SOD1 in both human disease (112) and mouse models (67,114–118).

Axonal degeneration, retarded axonal transport and accumulation of insoluble TAU have been observed in TAU\textsuperscript{P301L} (60), TAU\textsuperscript{R406W} (119), TAU\textsuperscript{P301S+S202V} (120), TAU\textsuperscript{K360I} transgenic mice (121) and in human neuroblastoma cells (122). However, neither deletion nor overexpression of human TAU isoforms impairs axonal transport, indicating a specific effect of mutated TAU protein (123).

Mutations in the p150 subunit of dynactin (DCTN1), a dynein activating macromolecular protein complex, have been reported in familial ALS, ALS with FTD (81,124), distal spinal and bulbar muscular atrophy (82,125) and Perry syndrome (126). The largest subunit, DCTN1, of the dynactin protein complex binds to both cytoplasmic dynein and microtubules. A recent study demonstrated that the C terminus of p150 directly binds to the N-terminal domain of TAU protein, enhancing the attachment of dynactin to microtubules (127). Interestingly, expression of another subunit of dynactin complex, dynactin 3 (p22/Dctn3, TAUmn\textsuperscript{+}; P < 0.001), which also binds directly to the p150 subunit, was significantly reduced in the TAU\textsuperscript{P301L} motor neurons (Supplementary Material, Table S2). Its biological function, chromosomal localization (9p13) and identification in TAU\textsuperscript{P301L} motor neurons made DCTN3 an attractive candidate for sequencing...
in FTD/ALS families linked to 9p21–13, as in an array-based gene discovery approach used previously (128,129). Although we have not identified mutations in the examined families, our study was limited by its small sample size (n = 3). A more comprehensive screening study is needed.

**MMP14, TIMP-2 and TIMP-3 dysregulation**

Protein expression of membrane-inserted matrix metalloproteinase 14 (Mmp14; TAUmn↓) was significantly decreased in both TauP301L mouse and SALS motor neurons (Table 1; Figs 4 and 5). Mmp14 is a member of a large family of matrix metalloproteinases (MMPs) whose activities are regulated by a specific protein family of tissue inhibitors of matrix metalloproteinases (TIMPs) [TIMPs (130,131)]. MMPs regulate remodeling of the extracellular matrix and thus play a key role in a wide range of physiological processes in the CNS, including growth, development and repair (132). Involvement of the MMP–TIMP system in the control and/or development of neurodegenerative diseases including ALS is being extensively studied (132) with a notable role of MMP2 and MMP9 in neural repair and as blood biomarkers of the disease progression (133–135). Mmp14 was expressed in the brain (136) but its dysregulation was not associated with early or late stages of motor neuron degeneration. Mmp14 was inhibited by TIMP2, 3 and 4 (137) and was linked to multiple sclerosis, malignant gliomas, stroke and Alzheimer’s disease (138). Two of the Mmp14 regulators, TIMP2 and TIMP3, were found in our mouse microarray study (Supplementary Material, Table S2). While activated at the transcriptional level in glia of SOD1G93A (Timp2) and TauP301L (Timp3) transgenic mice, their protein expression was significantly decreased in postmortem SALS spinal cord tissues (Fig. 6), in agreement with previously published microarray study of postmortem SALS tissues (46). Reduction of TIMP3 expression in the course of the disease is also consistent with a recent observation that the levels of TIMP3 are increased in degenerating motor neurons in SOD1G93A mice early in the disease course (8 weeks) and decreased at later stages of the disease proportionally to the remaining number of live motor neurons (139). These observations support the hypothesis that TIMP3 is a neuronal apoptotic protein whose early activation contributes directly to motor neuron death.

TIMP2 (SODgL↓) is most abundantly expressed in the adult CNS (140) where unlike TIMP3 it plays a neuroprotective function (141,142) through inhibition of matrix metalloproteinases (143). Increases in TIMP2 were observed in the CSF of patients with Alzheimer’s and Huntington’s diseases (141), and its elevation in presymptomatic SOD1G93A glia may be controlled by similar mechanisms.

**Peripheral blood transcriptome as a source of ALS clinical biomarkers and drug targets**

Recently several efforts were made toward the identification of early diagnostic and prognostic blood based markers of ALS (144). So far, somewhat inconsistent results point to the potential use of several blood biomarkers mainly applicable to the prediction of disease progression or therapeutic response. No clinical ALS biomarkers exist yet for early disease diagnosis that can be used prior to any notable clinical phenotype. This, in turn, prevents any attempts of early treatment using disease modifying drugs (144,145). Existing candidate ALS blood biomarkers range from single amino acid and protein to a combined panel of proteomic biomarkers. Surprisingly, significantly fewer attempts to identify markers based on the peripheral blood transcriptome were conducted despite an increasing volume of data supporting the notion that the peripheral blood transcriptome is a compelling source of biomarkers for neurological diseases. The ability of the blood transcriptome to respond specifically to the underlying pathogenic processes occurring in CNS is becoming evident with new data arising from the studies of patients with ischemic stroke (146,147), seizures (146), autism (148,149), multiple sclerosis (150), Tourette syndrome (151), Alzheimer’s disease (152) and other neurodegenerative disorders including ALS (153,154). Recent work by Lincecum et al. (153) convincingly demonstrates that genes whose expression levels are altered in SOD1G93A mice may also be found in the peripheral blood transcriptome of up to 65% of human ALS patients and specifically labels regulatory pathways identified in affected CNS tissues. However, most identified gene expression changes were associated with a fully developed disease phenotype and no studies of blood transcriptome in presymptomatic animals were reported raising an issue whether identified genes may be used as clinical biomarkers of the disease (Supplementary Material, Table S1). Here, for the first time, we report the identification of several ALS-relevant and presymptomatic blood gene expression changes in a well-established mouse model of familial ALS, SOD1G93A. Our studies demonstrated that at least 10% of the gene expression changes found in motor neurons and glia of presymptomatic SOD1G93A mice may also be detected in the blood (Table 3). Three of the 13 genes confirmed in blood—neurofilament heavy chain (Nefh; SODGl↓; SODb↓), peripherin (Prph; SODGl↓; SODb↓) and monoglyceride lipase (Mgl; SODGl↑; SODb↑) represent known functional contributors to ALS pathogenesis. Neurofilament heavy chain (Nefh; SODGl↓; SODb↓) and Peripherin (Prph; SODGl↓; SODb↓) are both intermediate filaments involved in the stabilization of motor neuron cytoskeleton whose mutations were associated with some familial ALS cases (61–63,65–70).

Intracellular accumulation of Nefh is one of the major pathological hallmarks of advanced ALS (155–157) and neurodegeneration in general (158). An increase in hyperphosphorylated Nefh in serum or CSF is also an important biomarker of axonal damage or a degree of neuronal loss after spinal cord injury or traumatic brain injury (159,160) and autoimmune encephalomyelitis (161). The expression of Nefh is regulated during CNS development and marks neuronal differentiation (162). Nefh is also expressed in blood and regulated independently from light and medium neurofilament subunits (Nefl and Nefm); however, molecular mechanisms of regulation remain unknown.

Prph (SODmn↓; SODb↓) is a type III intermediate filament predominantly expressed in neurons and capable of forming filamentous networks (157). Prph protein expression is upregulated after CNS injury and is attributed to regenerating neurons (reviewed in 157). Similar to TAU, peripherin is
capable of forming intraneuronal aggregates and may be neurotoxic if overexpressed (61,163). Pathological Prph abnormalities including mutations are associated with both FALS and SALS (62,164–166). Prph protein is upregulated in motor neurons of FALS and SOD1 \(^{G93A}\) mice (63). This increase of Prph correlated with disease onset, not being observed in motor neurons of presymptomatic mice (63). Given our data showing downregulated Prph gene expression in presymptomatic SOD1 \(^{G93A}\) mice and extensive motor neuron death in transgenic Prph mice overexpressing peripherin (61), it is tempting to speculate that the gradual Prph protein increase in the course of the disease is a neurotoxic factor contributing to eventual motor neuron death. Low transcriptional levels of both Nefh and Prph in motor neurons and blood of presymptomatic SOD1 \(^{G93A}\) mice point to a potential compensatory transcriptional downregulation in response to the excess of or accumulating corresponding protein products (Table 3; Supplementary Material, Table S2).

Monoglyceride lipase (Mgll; SODgl\(^{\uparrow}\), SODbh\(^{\uparrow}\)) is a key enzyme inactivating endogenous cannabinoids [ECs (167)] and one of the main drug targets for modulating EC system activity (168). Neuroprotective properties of ECs have extensively been studied in different neurodegenerative pathologies including ALS (168). ECs activation is a promising therapeutic approach already supported by multiple studies in ALS models (169,170). More specifically, it was demonstrated that chemical agonists of endocannabinoid receptors and elevated levels of ECs apply strong anti-inflammatory and neuroprotective effects, delaying disease progression in SOD1 \(^{G93A}\) mice (171–173). Activation of EC system reduces the release of glutamate from presynaptic terminals lowering the risk of neuronal death (168,170). Mgll is one of the two key enzymes that inactivate ECs and therefore its inhibitors selectively and locally enhancing EC signaling are attractive alternatives to the globally acting EC receptor agonists (174,175). In our study, Mgll gene expression was activated in SOD1 glia (Supplementary Material, Table S3) and was also increased significantly (\(P = 0.011\)) in the peripheral blood of presymptomatic SOD1 \(^{G93A}\) mice, suggesting that in addition to an important drug target it may also be a clinical biomarker for ALS.

In conclusion, we used large-scale comprehensive screens to identify pathogenic mechanisms in ALS, using SOD1 and TAU mouse models and human SALS tissues. We identified CNGA3, CRB1 and OTUB2 as putative common signaling factors in SOD1-, TAU- and SALS-related motor neuron degeneration. We identified CNGA3, CRB1, OTUB2, SLK, DDX58, RSPO2 and MMP14 as novel proteins specifically dysregulated in SALS. Using Chilibot and literature searches, we identified possible interactions between SOD1 and TAU that could explain the phenotypic similarities in these murine models, despite sharing only eight motor neuron and glia-expressed genes in common. In addition, analysis of SOD1 \(^{G93A}\) blood transcriptome identified significant changes in the expression of the 13 genes, also regulated in SOD1 \(^{G93A}\) motor neurons or glia, supporting their use as feasible clinical biomarkers. Although the pathogenesis of ALS remains to be elucidated, we have established an efficient approach for the discovery of factors relevant to a complex human neurological disease.

MATERIALS AND METHODS

Ethics statement

All animal experimentation was approved by the UCLA Animal Research Committee.

Mice

SOD1 \(^{G93A}\) mice, strain B6SJL-Tg(SOD1-G93A)1Gur (59), were obtained from The Jackson Laboratory (Bar Harbor, ME, USA). This strain expresses the transgene under the control of the endogenous human promotor (176). For all studies, female SOD1 \(^{G93A}\) transgenic and nontransgenic control littersmates were sacrificed at 8 weeks of age before the onset of the known disease pathology.

The TAU \(^{P301L}\) strain (Taconic), originally obtained from Michael Hutton’s group (60), was maintained by breeding hemizygous TAU \(^{P301L}\) mutant mice to the nontransgenic B6D2F1 strain. TAU \(^{P301L}\) mice express the longest 4R0N isoform of the most commonly found FTDP-17 TAU mutation (P301L) encoding for four-repeat tau without N-terminal inserts under the control of the mouse prion promoter (177,178). TAU \(^{P301L}\) mice demonstrate tau protein hyperphosphorylation and accumulation of neurofibrillar tangles accompanied by motor neuron degeneration in the spinal cord starting at 6 months, with more moderate pathology in the motor cortex. Motor deficits are detected in transgenic TAU \(^{P301L}\) hemizygous mice as early as 6.5 months (60). For this study, 16-week-old female presymptomatic TAU \(^{P301L}\) transgenic and nontransgenic control littersmates were used.

Breeding, including weaning at 3 weeks and genotyping, was performed according to the established protocols (20,60). Lumbar spinal cords from four SOD1 \(^{G93A}\) females and four TAU \(^{P301L}\) females, as well as from corresponding nontransgenic sex-matched littermate controls, were collected as described previously (3).

Laser capture microdissection

Rapidly harvested and frozen mouse lumbar spinal cords were cut into 12 \(\mu\)m transverse sections and placed onto non-charged glass slides, quickly stained with Cresyl violet and dehydrated according to the standard protocols [Arcturus, Molecular Devices, CA, USA (102)]. LCM of lumbar anterior horn motor neurons and surrounding glia was performed using the PixCell \(^{\text{TM}}\) II LCM System (Arcturus). For motor neuron collection, only the large cells in the ventral horns of the spinal cord stained with Cresyl violet were selected for microdissection. Regions with smaller punctuate staining of Cresyl violet were used for collection of surrounding glia (Supplementary Material, Fig. S3A). On average, 500 motor neuron bodies or surrounding glial cells from 20 tissue sections per animal were collected.

RNA isolation, amplification and labeling

RNA isolation from LCM-collected cells was performed using the Qiagen RNeasy Micro Kit with a slightly modified manufacturer’s protocol for optimal yield. Up to three cap linings containing the cells were collected and placed in 75 \(\mu\)l of
RNA lysis tissue buffer on ice until all cells were harvested. Up to four caps were processed through one column. The quality of the RNA was verified after every LCM procedure using the Agilent Bioanalyzer 2100 (Supplementary Material, Fig. S4). Isolated RNA consistently demonstrated high integrity and lack of degradation (Supplementary Material, Fig. S4). Only RNA with RNA integrity numbers > 7.5 were taken for further analysis. The amount of RNA isolated from 500 cells was sufficient for a two-round T7-based amplification and labeling protocol [Agilent Low RNA Input Linear Amplification Kit PLUS (6,102)] and consistently generated high-quality labeled cRNA (Supplementary Material, Fig. S4). Dye incorporation efficiency was checked with an ND-1000 spectrophotometer (NanoDrop Technologies, Rockland, DE, USA) and the Agilent Bioanalyzer 2100.

Whole-blood samples taken from mouse tail in 200 µl aliquots were processed for total RNA using the Whole-blood RNA MiniPrep Kit (Zymo Research, CA, USA). RNA concentration and quality were determined using the NanoDrop 2000 Spectrophotometer and the Agilent 2100 Bioanalyzer. One hundred nanograms of total RNA were used for one-round T7-based amplification and labeling (Agilent Low RNA Input Linear Amplification Kit PLUS).

Microarray hybridizations and data analysis

Agilent Whole Mouse Genome Microarray comparisons of motor neurons and surrounding glia were performed between transgenic (SOD1G93A and TAUP301L) and corresponding nontransgenic (control) littermate animals, producing four independent comparison groups: SOD1G93A motor neurons versus control motor neurons (SOD1mn), SOD1G93A motor neuron surrounding glia versus control glia (SOD1gl), TAUP301L motor neurons versus control motor neurons (TAUmn) and TAUP301L glia versus control glia (tauGl) (Fig. 1). Each comparison used four pairs of transgenic (SOD1G93A or TAUP301L) versus corresponding control littermate animals, producing four biological replicates. Each replicate was repeated twice with dye flip to correct for unequal dye incorporation rate. Therefore, eight microarray hybridizations were performed for each biological comparison, for a total of 32 microarray slides and generating four groups of differentially expressed genes in SOD1G93A motor neurons (SOD1mn), SOD1G93A motor neuron surrounding glial cells (SOD1gl), TAUP301L motor neurons (TAUmn) and TAUP301L surrounding glial cells (TAUgl). The direction of expression for each particular gene is indicated with up (†) or down (‡) arrows in the text. It should be noted that collected glial cells surrounding motor neurons also incorporated some of the neuronal processes.

Raw microarray data were acquired using the Agilent DNA Microarray scanner and processed with the accompanying Agilent Feature Extraction 10.5 Image Analysis software using default settings (see MIAME report). Normalized signal intensities were used to identify gene expression changes in SOD1G93A and TAUP301L motor neurons and surrounding glial cells, generating four partially overlapping sets of data. For the identification of differential expression, the genes were required to pass two conservative criteria: a ratio beyond the 99.5% confidence interval observed in homotypic comparisons (3), which corresponded to an ~1.5-fold expression change and a paired t-test (P < 0.01) computed using 100 permutations of the data for each gene. Correction for multiple comparisons was performed using the adjusted Bonferroni test. The analysis was performed in the TM4: Microarray Software Suite (http://www.tm4.org/ (179)).

To screen for gene expression changes in the peripheral blood of presymptomatic (10 weeks) SOD1G93A transgenic mice, a custom SOD-microarray was designed using Agilent eArray technology (https://earray.chem.agilent.com) with oligonucleotide probes corresponding to the genes found in the whole-genome microarray screen of motor neurons and glia (see MIAME report). To increase the chances of identifying differentially regulated genes in the blood, more relaxed criteria were applied for preselection of genes for the SOD-array. Vulnerable (‘regulated’) genes were identified using unpaired t-test assuming equal variances (P < 0.05) between experimental samples (SOD1G93A) and nontransgenic controls. This resulted in 1449 genes on the SOD-array. Identification of genes regulated in SOD1G93A peripheral blood (SODb) was done according to the same protocol as in the initial whole-genome microarray screen.

GO and pathway analysis

To assess the relevance of the identified gene expression changes, searches for GO-based, overrepresented functional groups and literature-based functional networks were performed using the Database for Annotation, Visualization, and Integrated Discovery [DAVID; http://david.abcc.ncifcrf.gov/ (180,181)], Pubmatrix [http://pubmatrix.grc.nia.nih.gov; (182) and Chilibot [http://www.chilibot.net (183)] online bioinformatics tools. Identification of enriched categories and biological processes among regulated genes was performed using DAVID Functional Annotation Chart. To avoid the errors due to duplicated genes the Fisher exact statistics was calculated based on corresponding DAVID gene IDs. The significance of enrichment or EASE score is a modified Fisher exact P-value with Benjamini correction for multiple comparisons. Only categories with EASE score < 0.05 were considered significant.

Semi-quantitative RT–PCR

To provide technical confirmation of microarray data, selected genes were subjected to semi-quantitative RT–PCR analyses. Assays were performed using T7-amplified cRNA templates generated from total RNA isolated from LCM motor neurons of SOD1, TAU and control animals. Initially 100 ng of cRNA were converted to cDNA using polyT primers. RT–PCRs were performed using the ImProm-II Reverse Transcription System (Promega) according to manufacturer’s recommendations. For validation of microarray-identified glia-specific expression changes, total RNA isolated from spinal cord tissues (RNeasy-Mini Kit, Qiagen) of SOD1G93A, TAUP301L and control mice were used as templates. Amplification primers were designed to span exon–exon junctions to eliminate potential genomic DNA amplification. Prior to semi-quantitative RT–PCRs, all primers were tested and conditions were optimized to ensure a single product band.
Primers for mouse ribosomal 18S RNA were used in the same RT–PCR for normalization. The sequences of primers and the reaction conditions are shown in Supplementary Material, Table S6. RT–PCR products were analyzed in 6% PAAG stained with SYBR Green I dye and scanned using the Typhoon 9419 system (Amersham Biosciences). Each experiment included four independent biological samples and was performed at least twice. The significance of gene expression change was assessed using Student’s t-tests.

**Western blot analysis and antibodies**

Dissected human postmortem frozen spinal cord anterior horn tissue samples were homogenized by sonication in hypotonic buffer (10 mM Tris–HCl, 10 mM KCl, 0.1 mM EDTA, 0.1% Triton X-100, pH 8.0) with protease inhibitor cocktail (Sigma). Protein extracts were centrifuged for 20 min at 10 000g at 4°C. Supernatants were decanted and protein concentrations in the crude extract were determined using Coomassie Protein Assay Reagent (Pierce). Equal amounts of protein (50–70 μg) were loaded and separated by SDS–PAGE using double 6–12% Tris–HCl gels. Proteins were transferred to a polyvinylidene fluoride membrane for 3–14 h at 4°C, incubated in blocking buffer [5% nonfat milk in tris-buffered saline and tween 20 (TBST)] for 1 h, and primary antibodies (Supplementary Material, Table S3) were applied overnight at 4°C (three times for 5 min), membranes were incubated with peroxidase-conjugated secondary antibodies for 1 h at room temperature and washed with TBST (three times for 5 min). Antibody detection was performed using the ECL kit (GE). To quantify the resulting bands, the blots were scanned on a Typhoon 9419 system (Amersham Bioscience) at an appropriate excitation and emission wavelengths (ImageQuant 5.2). Primary antibodies included rabbit polyclonal anti-DDX58 (1:100, LifeSpan), rabbit monoclonal anti-MMP-14 (1:100, Gentex) and mouse monoclonal anti-Cnga3 (1:50, Santa Cruz), and were kept at −80°C. Immunohistochemistry was prepared according to standard protocols (Vector). Primary antibodies used were goat polyclonal anti-Dusp7 (1:500, NovusBio), rabbit polyclonal anti-Timp2 (1:1000, Sigma), rabbit polyclonal anti-Timp3 (1:2000, Biomol International) and R-spondin 2 (1:1000, PTG, Supplementary Material, Table S3). Data were averaged from at least three separate experiments. The statistical significance was estimated by Student’s t-tests.

**Mouse immunohistochemistry**

Mouse lumbar spinal cord regions were cut into 12 μm sections and placed onto charged glass slides (at least four sections per slide) and air dried for 20 min. For long-term storage, slides were kept at −80°C. Immunohistochemistry was performed according to standard protocols (Vector). Primary antibodies used were goat polyclonal anti-Cnga3 (1:50; Santa Cruz), rabbit monoclonal anti-MMP-14 (1:100; Gentex) and mouse monoclonal anti-DDX58 antibody (LifeSpan; Supplementary Material, Table S3).

**Human subjects, DNA isolation and sequencing**

Blood samples were collected from 12 members of 3 ALS/FTD families of Canadian descent that were linked to chromosome arm 9p [Que 1, Que 23 and Fr 104 (55)]. Of the 12 samples, 6 were unaffected and 6 were affected or at-risk for FTD/ALS. Each family was represented by at least one member unaffected with ALS and one member affected with ALS. Each subject was clinically assessed, and signed, informed consent was obtained. For controls, DNA from the blood of healthy individuals was used (Invitrogen), isolated by standard methods (DNeasy kit, Qiagen). A total of 21 primer pairs for CCL19 and DCTN3 were designed and synthesized (Supplementary Material, Table S7). Seven primer pairs were synthesized to amplify four exons and UTRs of CCL19. Fourteen primer pairs were synthesized to amplify seven exons and UTRs of DCTN3. DNA from each sample was amplified and the product was purified using a Qiagen DNA purification kit. The purified products were checked on 1% agarose gels and sequenced using ABI sequencers. Non-coding 5′ and 3′ UTRs and coding regions with adjacent intronic sequences (at least 200 bp on each end) of identified candidate genes were analyzed for sequence changes (Supplementary Material, Table S5).

**Human postmortem tissues and ALS TMA (ALS–TMA) construction**

Postmortem paraffin-embedded brain and spinal cord tissue blocks from SALS (19 subjects) and normal controls (6 subjects), provided by the UCLA Department of Pathology and the National Neurological AIDS BANK (NNAB), were used for TMA construction at the UCLA Tissue Array core facility (http://www.genetics.ucla.edu/tissuearray). Using spinal cord material and control CNS regions from ALS and control subjects, 719 cylindrical (0.6 mm diameter) cores of formalin-fixed and paraffin-embedded tissues were densely and precisely assembled into two paraffin blocks (Supplementary Material, Fig. S1). Each area of interest (for cores) within hematoxylin and eosin (H&E) stained donor tissue was identified under a light microscope. In order to minimize the incidence of microfractures in the recipient block, each block was softened by incubation at 37°C for 5 min before extraction of recipient cores. Careful preparation and microtome cutting of TMA blocks resulted in ~90% of tissue cores being available for analysis. Origins of donor cores were recorded on digital images of H&E-stained tissue sections from the original donor blocks. The donor blocks represented: 2–4 cores from each of two brain regions (grey and white matter) and 2–3 cores from each of the three spinal cord regions— anterior horn, posterior horn and lateral spinal tracts. From the assembled blocks, over 400 serial 8 μm thick TMAs were produced and placed on individual glass microscope slides. An adhesive-coated tape system (Intrumedic, NJ, USA) was used for maintaining optimal array formatting (ImageJ; Fire, USA) was used for maintaining optimal array formatting when transferring cut tissue ribbons to glass slides. To ensure the quality of the resulting TMAs, morphological bright field examination, a panel of traditional histological stains (e.g. H&E, Kluver), as well as neuron and glia staining with NeuN and GFAP, respectively, were performed (not shown).

**TMA immunohistochemistry**

After deparaffinization (3), antigen retrieval was performed by microwave boiling in 0.01 M sodium citrate buffer (pH 6.0) for
10 min, followed by cooling down for 20 min at room temperature (RT). Following three 5 min washes with PBS, tissue cores were covered with blocking buffer (5% BSA, 5% normal serum in PBS) and incubated for 1 h at RT. Incubation with primary antibodies was performed overnight at 4°C, followed by washing with PBS (three times for 5 min). Secondary biotinylated antibodies (Vector) were applied for 30 min at RT, followed by incubation with the avidin–biotin complex (ABC ‘Elite’, Vector lab) for 30 min at RT. Slides were then rinsed three times for 5 min with PBS and incubated with 3,3′-diaminobenzidine (DAB) solution (Vector) until suitable staining developed (generally 2–5 min) and counterstained with hematoxylin (Vector). Primary antibodies included mouse monoclonal anti-CCL-19 (1:40; R&D), goat polyclonal anti-ChAT (1:100; Chemicon), goat polyclonal anti-Cnga3 (1:50; Santa Cruz), goat polyclonal anti-Crb1 (1:50; Santa Cruz), rabbit polyclonal anti-Dcx58 (RIG-1, 1:100; LifeSpan), mouse monoclonal anti-DHC complex (1:10; Sigma), rabbit polyclonal anti-DHC complex (1:100; Santa Cruz), rabbit monoclonal anti-MMP-14 (1:100; Gentex), mouse monoclonal anti-neuronal nuclei (NeuN; 1:100; Chemicon), mouse anti-Otub2 serum (1:50; Novus Biol, Inc.), mouse monoclonal anti-S100a9 (CalB; 1:250; Acris) and mouse monoclonal anti-SKL antibody (1:50; Novus Biol, Inc.; Supplementary Material, Table S3).

TMA data analysis
Cores that were severely damaged during tissue processing and staining were excluded from the analysis. The histology of all cores was verified on H&E-stained sections to confirm that a region of interest from a donor block appeared in a given core. Analysis of motor neuron staining was performed in TMA slide pairs (control/experiment) representing adjacent serial 8 μm tissue sections. Therefore, immunostaining for different antibodies for motor neuron-specific expression (NeuN staining) and controls could be examined in the same cells. Immunostaining with omission of primary antibodies were used as negative controls.

Two types of analyses were applied to estimate the relative protein expression in SALS versus control tissues. First, the analysis focused on the number of cells expressing the specific protein in a given sample or subject. The immunoreactivity was judged as positive, regardless of staining intensity, when the signal in a given cell was higher than the background. The first analysis (TMA1) also took into account the average protein expression, which was calculated as the mean for all motor neurons in a given tissue core minus the mean background. The resulting average protein expression values for each core were grouped according to phenotype (control and SALS), and the difference was estimated using a standard t-test. In the second analysis (TMA2), the total number of motor neurons identified based on morphology and NeuN staining was divided by the total number of cells positive for a protein of interest, and a specific protein expression index (PEI) for a given tissue core was derived. The average PEI was calculated for each tissue type and individual. Statistical significance was based on a Student’s t-test (P < 0.01). To ensure that the changes in protein expression were not artifactual, each immunostaining experiment was performed at least twice.

SUPPLEMENTARY MATERIAL
Supplementary Material is available at HMG online.

ACKNOWLEDGEMENTS
The authors thank Dr Patricia Dickson and Dr Henry Lin for critical reading of the manuscript. We thank Dr Elise Singer from the National Neurological AIDS Bank, and Dr Wallace Tourtellotte from the Human Brain and Spinal Fluid Resource Center Neurology Service for providing valuable human postmortem tissue.

Conflict of Interest statement. None declared.

FUNDING
This work was supported by grants from the Muscular Dystrophy Association (S.L.K., M.W.-P.), the Amyotrophic Lateral Sclerosis Association (S.L.K., M.W.-P.), the Judith and Jean Pape Adams Charitable Foundation (S.L.K., M.W.-P.), RGK Foundation (S.L.K., M.W.-P.), NARSAD (S.L.K.), Alzheimer Association (S.L.K.), and CurePSP Foundation (S.L.K.).

REFERENCES


