Human Molecular Genetics, 2019, Vol. 28, No. 16 2799–2810

doi: 10.1093/hmg/ddz104
Advance Access Publication Date: 20 May 2019
Bioinformatics Article

BIOINFORMATICS ARTICLE

Motor neurons from ALS patients with mutations in *C9ORF72* and *SOD1* exhibit distinct transcriptional landscapes

Ching-On Wong and Kartik Venkatachalam*,†

Department of Integrative Biology and Pharmacology, McGovern Medical School at the University of Texas Health Sciences Center (UTHealth), Houston, TX 77030, USA and Graduate Program in Biochemistry and Cell Biology, MD Anderson Cancer Center and UTHealth Graduate School of Biomedical Sciences, Houston, TX 77030, USA

*To whom correspondence should be addressed at: Department of Integrative Biology and Pharmacology, McGovern Medical School at the University of Texas Health Sciences Center (UTHealth), 6431 Fannin St, MSB 4.214, Houston TX 77030, USA. Tel: +713 5007504; Fax: +713 5007456; Email: kartik.venkatachalam@uth.tmc.edu

Abstract

Amyotrophic lateral sclerosis (ALS) is a progressive motor neuron disease that culminates in paralysis and death. Here, we present our analyses of publicly available multiOMIC data sets generated using motor neurons from ALS patients and control cohorts. Functional annotation of differentially expressed genes in induced pluripotent stem cell (iPSC)-derived motor neurons generated from patients with mutations in *C9ORF72* (C9-ALS) suggests elevated expression of genes that pertain to extracellular matrix (ECM) and cell adhesion, inflammation and TGFβ targets. On the other end of the continuum, we detected diminished expression of genes repressed by quiescence-promoting E2F4/DREAM complex. Proteins whose abundance was significantly altered in C9-ALS neurons faithfully recapitulated the transcriptional aberrations. Importantly, patterns of gene expression in spinal motor neurons dissected from C9-ALS or sporadic ALS patients were highly concordant with each other and with the C9-ALS iPSC neurons. In contrast, motor neurons from patients with mutations in *SOD1* exhibited dramatically different signatures. Elevated expression of gene sets such as ECM and cell adhesion genes occurs in C9 and sporadic ALS but not SOD1-ALS. These analyses indicate that despite the similarities in outward manifestations, transcriptional and proteomic signatures in ALS motor neurons can vary significantly depending on the identity of the causal mutations.

Introduction

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease that affects both the upper and lower motor neurons leading to paralysis and death due to respiratory failure (1,2). The etiology of ALS is incredibly complex and multifactorial, which result in aberrations in a variety of pathways. Pathways affected in ALS include mRNA processing and trafficking, protein homeostasis, nucleocytoplasmic transport, endoplasmic reticulum stress and mitochondrial function (1,2). Consistent with the myriad biological processes implicated in ALS, >20 genes are associated with the ALS and the ALS–frontotemporal dementia (ALS–FTD) continuum (1,2). Hexanucleotide repeat expansions in *C9ORF72* are one

†Kartik Venkatachalam, http://orcid.org/0000-0002-3055-9265
Received: November 5, 2018. **Revised:** April 23, 2019. **Accepted:** May 13, 2019

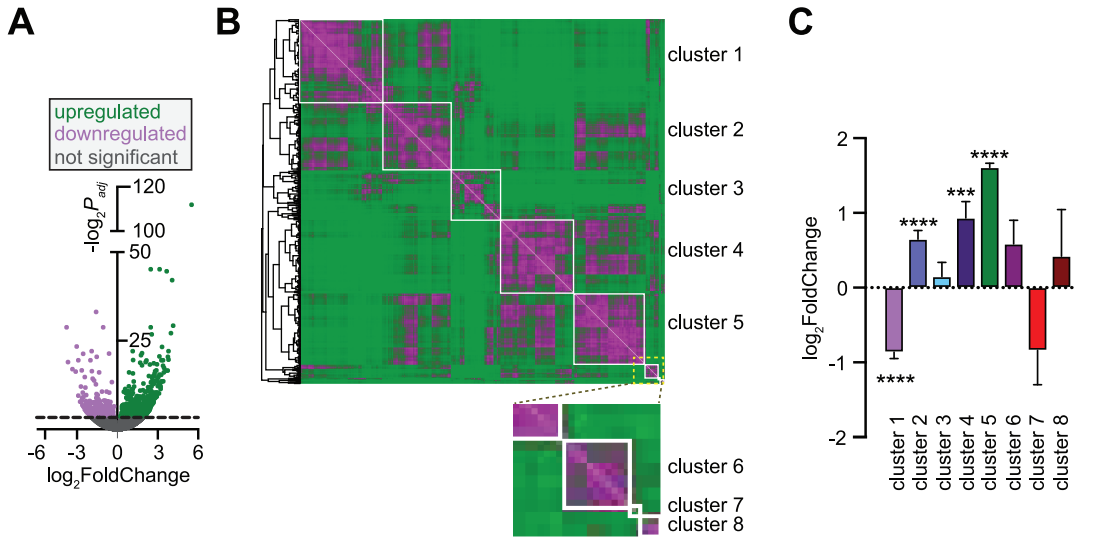

Figure 1. Analysis of differential gene expression in iPSC-derived motor neurons generated from patients with C9-ALS. (**A**) Volcano plot showing log values of the fold change in the expression of protein-coding genes in C9-ALS motor neurons plotted against negative log value of the adjusted *P*-values determined by DESeq2. Horizontal *dashed line* indicates the cutoff for significance ($P_{adj} = 0.1$). Up- and downregulated genes are represented by *green* and *magenta* circles, respectively. (**B**) Heatmap showing hierarchical clustering of the correlations of gene expression. A total of 860 DEGs are separated into 8 clusters as indicated. (**C**) Fold change of gene expression for each cluster in C9-ALS motor neurons. Values shown are mean ± standard error of the mean of the relative change of all genes within the same cluster. *Asterisks* indicate gene clusters that are significantly up- or downregulated ($P < 0.0001$, Wilcoxon signed-rank test against a hypothetical mean of 0 for clusters 1, 4 and 5 and one-sample *t*-test against a hypothetical mean of 0 for cluster 2).

of the most commonly occurring mutations in both the familial (fALS) and sporadic (sALS) forms of the disease (3–8). A seminal advance in delineating the relationship between patient-specific mutations and cellular alterations in motor neurons was generation of induced pluripotent stem cell (iPSC)-derived motor neurons from ALS patients (9,10). This technology has enabled detailed analyses of cell biological, genetic, transcriptional, epigenetic and proteomic changes in ALS and provided unprecedented insights into genetic disturbances that correlate with disease phenotypes (10,11). Indeed, comparative analyses using iPSC-derived motor neurons probing the overlap between fALS and sALS continue to uncover both common and distinct biological mechanisms driving disease pathology (10–15).

These studies have revealed that the underlying perturbations in ALS are functions of the causal molecular lesions. For instance, while defective RNA processing has been implicated in ALS–FTD caused by mutations in *C9ORF72*, *TARDBP* and *FUS* (1,16), ALS patients with dominant mutations in *SOD1* do not exhibit similar defects in RNA processing (2,16–18). These types of distinctions underscore the notion that ALS is actually a spectrum of diseases that exhibit varying extents of genotypic and phenotypic overlap. A natural extension of this understanding is that development of effective therapeutic options for ALS is predicated upon appropriate stratification of patients on the basis of genetic and cell biological features. Here, we asked whether cases of ALS caused by different mutations might be stratified on the basis of alterations in transcriptional landscape. We reasoned that identification of links between the causal mutations and transcriptional and proteomic features of the disease could enable development of therapeutic strategies that are most appropriate to mitigate pathology. Our bioinformatic analyses revealed distinct clusters of differentially expressed genes (DEGs) in cases of ALS caused by mutations in *C9ORF72* versus *SOD1*. Our findings pave the way for future studies that develop therapeutic options that selectively target individual-specific perturbations.

Results

DEGs in iPSC-derived motor neurons generated from C9-ALS patients

We obtained transcriptomic data from NeuroLINCS (http://neurolincs.org)—a part of the National Institutes of Health-Library of Integrated Network-Base Cellular Signatures (NIH-LINCS) program that serves as a repository of multiOMIC data sets generated from iPSC-derived motor neurons isolated from ALS patients with *C9ORF72* mutations (C9-ALS) (11), spinomuscular atrophy (SMA) patients and unaffected controls. Analysis of protein-coding mRNAs using DESeq2 [False Discovery Rate (*FDR*) < 0.1] (19,20) revealed 860 DEGs in C9-ALS motor neurons, with 560 genes that were significantly upregulated and 300 genes that were significantly repressed (Fig. 1A and Supplementary Material, Table S1A). To glean insights into regulatory networks that sculpt this signature, we applied unsupervised hierarchical clustering to the coefficients of pairwise correlations of gene expression. This approach generated a matrix comprised of eight separate clusters (Fig. 1B and Supplementary Material, Table S1B). Assessment of the overall differential expression associated with these clusters showed that cluster 1 was significantly repressed in C9-ALS motor neurons, whereas clusters 2, 4 and 5 were significantly upregulated (Fig. 1C).

Next, we used gene set enrichment analysis (GSEA) (21,22) to classify the C9-ALS DEGs into functional categories. We ranked DEGs on the basis of differential expression in C9-ALS and compared the ranked list with the curated (GSEA: C2) and gene ontology (GSEA: C5) gene sets available at MSigDB (*FDR* $q < 0.1$). Whereas many gene sets belonged to cancer-related

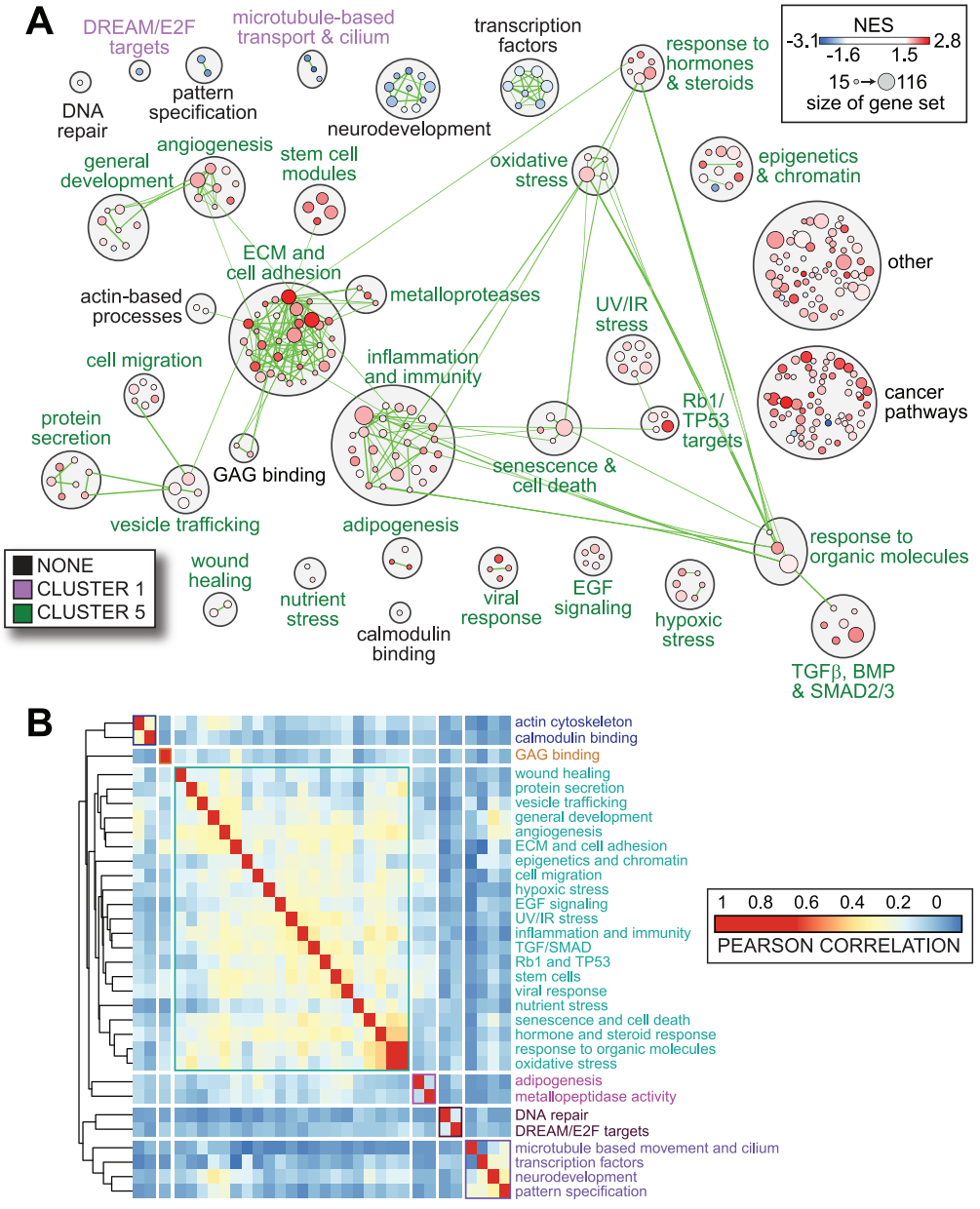

Figure 2. Functional annotation of DEGs in C9-ALS motor neurons. (**A**) Pathway enrichment analysis of DEGs in C9-ALS motor neurons. DEGs were ranked and subjected to unbiased GSEA. Gene sets with FDR $q < 0.1$ are shown as nodes. Node color represents NES from GSEA. Node size represents number of genes. Gene sets that belong to cluster 1 and cluster 5 in Figure 1B are labeled in *magenta* and *green*, respectively. Cluster attribution was by Fisher's exact test followed by Benjamini–Hochberg test using all the individual *P*-values (*FDR* < 0.05). (**B**) Heatmap showing hierarchical clustering of the Pearson correlations of the indicated modules.

pathways, others could not be grouped on the basis of ontology and were thus classified under the 'other' gene set (Fig. 2A and Supplementary Material, Table S2). A larger number of gene sets exhibited normalized enrichment score (NES) > 0 (Fig. 2A and Supplementary Material, Table S2), which reflected the overrepresentation of upregulated DEGs in the C9-ALS motor neurons (Fig. 1A and Supplementary Material, Table S1A). Two of the largest modules with NES > 0 were composed of gene sets pertinent to extracellular matrix (ECM) and cell adhesion and inflammation and immunity (Fig. 2A and Supplementary Material, Table S2). Other upregulated modules included metalloproteases and positive regulators of angiogenesis and wound healing. Various stress response pathways (oxidative, hypoxic, UV/IR and nutrient) were concomitantly potentiated in C9-ALS (Fig. 2A and Supplementary Material, Table S2). Potential involvement of Rb1 and TP53 in C9-ALS neurons, which was evident by elevated expression of their transcriptional targets, is consistent with roles of these proteins in neuronal cell death (23–25). Indeed, modules related to senescence and cell death were also activated in C9-ALS motor neurons (Fig. 2A and Supplementary Material, Table S2). Modules with NES < 0 were composed of the DREAM/E2F targets, neurodevelopment, DNA repair, pattern specification, microtubule-based transport and general transcription (Fig. 2A and Supplementary Table 2). Comparison with clusters of coregulated DEGs (Fig. 1B) showed that most modules were composed of DEGs that belonged to either cluster 1 or 5 (Fig. 2A—labeled in magenta and green, respectively). Gene sets mapped to cluster 1 were downregulated

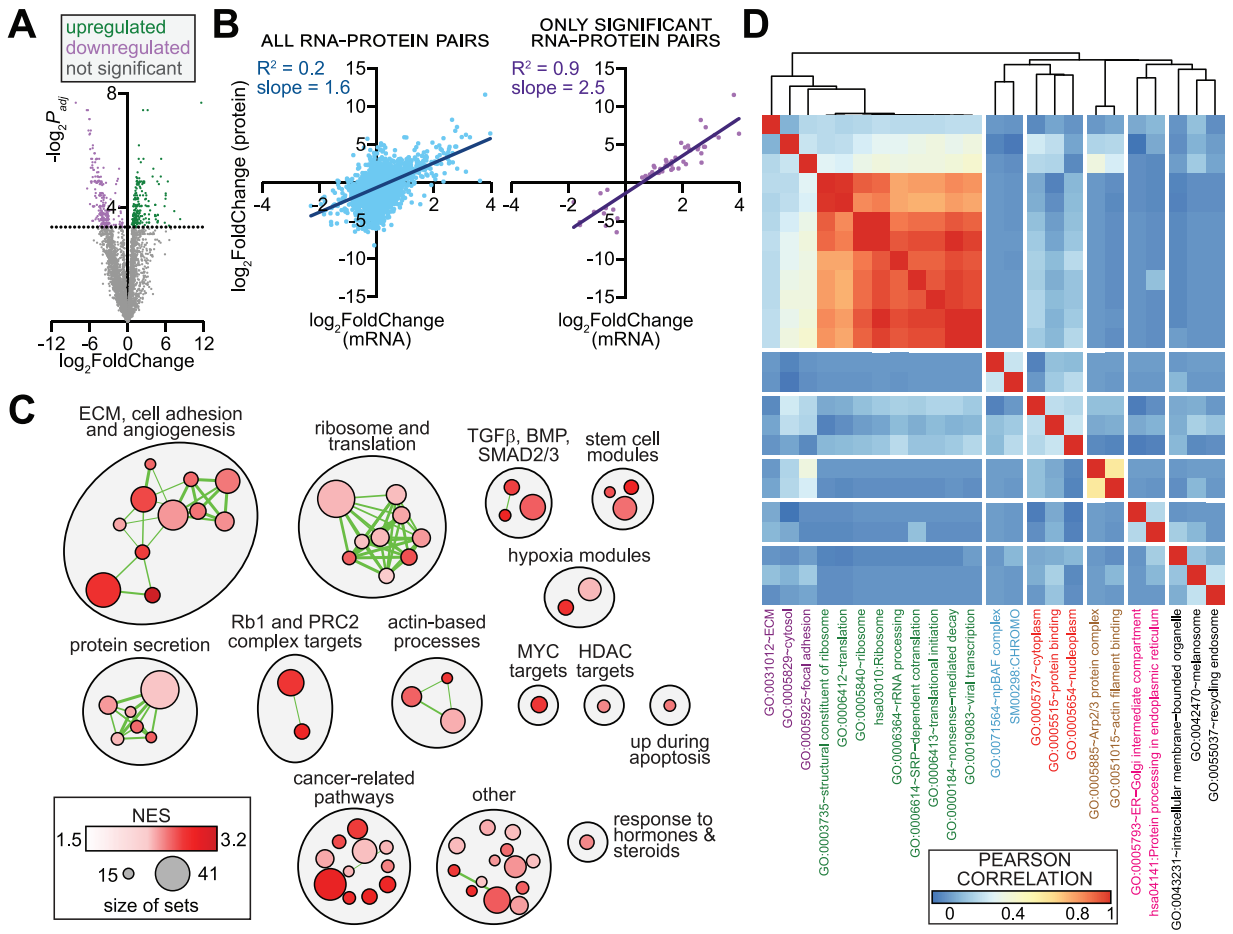

Figure 3. Concordance between transcriptomic and proteomic alterations in C9-ALS motor neurons. (**A**) Volcano plot showing log values of the fold change in DEPs in C9-ALS motor neurons plotted against negative log value of the adjusted *P*-values. Horizontal *dashed line* indicates the cutoff for significance ($P_{adj} = 0.1$). Proteins with increased and decreased abundance are represented by *green* and *magenta* circles, respectively. (**B**) Scatter plots showing relative changes in mRNA and protein abundance in C9-ALS motor neurons. Parameters of the linear fit are shown. (**C**) Functional annotation of upregulated proteins in C9-ALS motor neurons. All differentially abundant proteins were subjected to GSEA. Only sets with FDR $q < 0.05$ are shown as nodes. Node color represents NES from GSEA. Node size represents number of proteins that make up that set. (**D**) Heatmap showing hierarchical clustering of the correlations between the ontological categories. Differentially expressed proteins were assigned to indicated ontological categories using DAVID (Benjamini *FDR* < 0.05). Pairwise correlation between the various ontological categories was subjected to hierarchical clustering.

(NES < 0), whereas modules gene sets mapped to cluster 5 were upregulated (NES > 0). The fact that modules could be mapped to clusters of correlated genes suggests the involvement of higher-order regulatory pathways that coordinate the appearance of functional modules.

Analyses of genes that overlap between modules using hierarchical clustering revealed distinct clades (Fig. 2B). The largest of these clades comprised modules related to stress response, ECM and cell adhesion, inflammation and immune response, TGFβ and SMAD2/3 targets, Rb1 and TP53 targets, etc. (Fig. 2B). Smaller clades were composed of sets related to actin cytoskeleton, angiogenesis and metallopeptidases, DNA repair and DREAM/E2F target genes, microtubule-based movement, transcription factors, neuronal development and pattern specification (Fig. 2B).

Proteomic alterations in C9-ALS motor neurons mirror transcriptomic disturbances

We used the NeuroLINCS proteomics data set to assess whether the aforementioned alterations in mRNA in C9-ALS led to corresponding changes in protein abundance. We found that C9-ALS motor neurons exhibit 427 differentially enriched proteins (DEPs) consisted of 195 proteins with increased abundance and 232 proteins with lower abundance (Fig. 3A and Supplementary Material, Table S3A). Although a scatter plot of global changes in mRNA versus corresponding protein was a poor fit to a linear correlation, mRNA–protein pairs whose levels were significantly altered in C9-ALS exhibited a strong linear relationship (Fig. 3B). These data demonstrate high correlation between significantly altered mRNA and protein species in C9-ALS motor neurons. GSEA using gene names corresponding to the differentially abundant proteins allowed us to identify enriched sets that also appeared in the transcriptome including ECM, cell adhesion and angiogenesis and TGFβ/SMAD targets (Fig. 3C and Supplementary Material Table S3B). Although gene sets related to inflammation and immunity were observed in high density in the transcriptomic data (Fig. 2A), this module was not identified in the DEPs (Fig. 3C). This finding could reflect the simple fact that cytokines are released from cells, thus making them 'invisible' to proteomic analyses. On the other hand, DEPs in C9-ALS motor neurons showed an enrichment for proteins

involved in ribosome processing and mRNA translation (Fig. 3C), which was not reflected in the transcriptome. These data could indicate subthreshold increases at the mRNA levels that are sufficient to elicit enduring elevation in protein levels. Alternatively, our analyses could suggest diminished degradation of ribosomal proteins via autophagy (ribophagy) and the ubiquitin–proteasome system (26–28). Another intriguing aspect of the proteomic data was that despite the repression of several gene sets at the transcriptomic level, we could not identify repressed protein modules. In other words, proteins whose abundance was decreased in C9-ALS could not be sorted into clear functional categories. Thus, modules that were significantly repressed at the transcriptomic level were not similarly diminished in the proteome. This bias was not a consequence of protein abundance since the overall numbers of proteins whose abundance was increased versus decreased were very similar in the C9 motor neurons.

As an alternative to GSEA, we annotated the differentially abundant proteins using DAVID (Benjamini *FDR* < 0.05) (29,30). These data showed enrichment of the following ontological modules: ECM, focal adhesion, ribosomal processing and mRNA translation, actin cytoskeleton and protein/vesicular trafficking pathways (Fig. 3D and Supplementary Material, Table S3C). Overall, these data demonstrate appreciable concordance between the transcriptomic and proteomic data sets in C9-ALS motor neurons, especially for proteins related to ECM, cell adhesion and regulation of actin filaments. Retention of the ontological features of the C9-ALS transcriptome at the proteomic level lends credence to the notion that perturbations in gene expression may sculpt the cell biological features of C9-ALS motor neurons.

C9-ALS-associated transcriptional signatures are highly concordant in iPSC-derived and spinal motor neurons

We asked whether the transcriptional characteristics of iPSC-derived motor neurons generated from C9-ALS patients were also observed in spinal motor neurons dissected postmortem from patients with the disease. The rationale for a direct comparison was 3-fold. First, a similar signature in an independently generated data set would validate the conclusions generated using iPSC-derived motor neurons. Second, the appearance of a similar signature in spinal motor neurons isolated from C9-ALS patients would indicate that the transcriptional alterations identified in iPSC neurons are not an idiosyncratic feature of the process associated with the generation of iPSC-derived motor neurons. Third, comparison of iPSC and spinal motor neurons would allow us to assess whether the native niche of the motor neurons permits the transcriptional signature evident in iPSC-derived motor neurons. Keeping these factors in mind, we examined the transcriptional signature in spinal motor neurons microdissected postmortem from C9-ALS patients (15). Although differential expression in this data set was assessed using microarray rather than RNA-Seq (15), the overall transcriptional signatures were highly concordant between spinal and iPSC-derived motor neurons from C9-ALS patients (Fig. 4A and Supplementary Material, Table S4A). As was the case in iPSC neurons, gene sets that were significantly enriched in spinal motor neurons included ECM/cell adhesion, inflammation and immunity, stress response and angiogenesis (all with NES > 0) (Fig. 4A and Supplementary Material, Table S4A). The spinal motor neurons showed additional enrichment for gene sets related to activation of cell death, which was not evident in iPSC-derived motor neurons. Interestingly, gene sets that characterize Alzheimer's disease (AD) were also significantly enriched in C9-ALS spinal motor neurons with genes that were up- or down-regulated in AD being similarly altered in C9-ALS (Fig. 4A and Supplementary Material, Table S4A).

Some, although not all, patients with sALS exhibit features reminiscent of those observed in C9-ALS (5–8,16,31). Thus, we asked whether we might be able to detect similar transcriptomic alterations in sALS motor neurons. To answer this question, we examined an RNA-Seq data set generated from spinal motor neurons laser-microdissected postmortem from 12 individuals suffering from sALS and 9 unaffected controls (32). First, we examined whether DEGs in the C9-ALS and sALS data sets overlap using a simple contingency test. Briefly, we used the Fisher's exact test to determine the significance of co-occurrence between two parameters of the data sets—**(1)** the number of DEGs that occur in **both** data sets and **(2)** the number of DEGs that are similarly up- or downregulated in **both** data sets. Odds ratio of this contingency test tells us whether statistical significance relates to co-occurrence or mutual exclusion (>1 or <1, respectively). DEGs in spinal motor neurons from C9-ALS and sALS not only overlapped, the direction of differential expression was the same ('co-occurrence') (Fig. 4B). Similarly, we observed significant co-occurrence between DEGs in iPSC-derived motor neurons from C9-ALS and spinal motor neurons from sALS (Fig. 4C). These data are consistent with the previous descriptions that have probed the extent of transcriptional overlap between C9-ALS and sALS (33).

Even at the ontological level, DEGs in sALS were similar to those in C9-ALS. GSEA performed on sALS DEGs showed enrichment of several modules that were also observed in C9-ALS (Fig 4D and Supplementary Material, Table S4B). For instance, modules with overall NES > 0 included inflammation and immunity, ECM/cell adhesion and TGFβ/SMAD targets. Other upregulated modules related to cell migration/chemotaxis, endocytosis and vesicle trafficking, etc. On the other end of the gene expression spectrum, modules related to RNA processing, ribosomal biogenesis and protein translation and mitochondrial functions were repressed in sALS (Fig. 4D and Supplementary Material, Table S4B). Taken together, our meta-analyses point to significant transcriptional overlap in C9-ALS and sALS, especially in the case of upregulated genes that constitute functional modules related to ECM, cell adhesion and inflammation.

SOD1 mutations induce transcriptional signatures that are distinct from that observed in C9-ALS

Although hexanucleotide repeat expansion in C9ORF72 or *SOD1* mutations represent the most common genetic lesions associated with ALS, these mutations fall on either end of a broad genotype–phenotype spectrum (2). As an example, defects in RNA processing and nucleocytoplasmic transport are common occurrences in C9-ALS, but human or mouse motor neurons carrying *SOD1* mutations exhibit neither RNA processing defects nor alterations in nucleocytoplasmic transport (2,16–18). Intrigued by these differences, we asked whether the transcriptional signature in C9-ALS would appear in SOD1-ALS. To answer this question, we examined RNA-Seq data obtained from iPSC-derived human motor neurons bearing *SOD1* mutations and control cells in which the *SOD1* mutations had been genetically corrected (10,34). The results of GSEA using DEGs ranked on the basis of differential expression in $SOD1^{A272C}$ neurons (34) showed a signature that was the opposite of what was found in C9- or sALS (Fig. 5A and Supplementary Material, Table S5). In $SOD1^{A272C}$ neurons, modules related to ECM/cell adhesion, TGFβ/SMAD targets, actin-based processes and MAPK signaling were composed

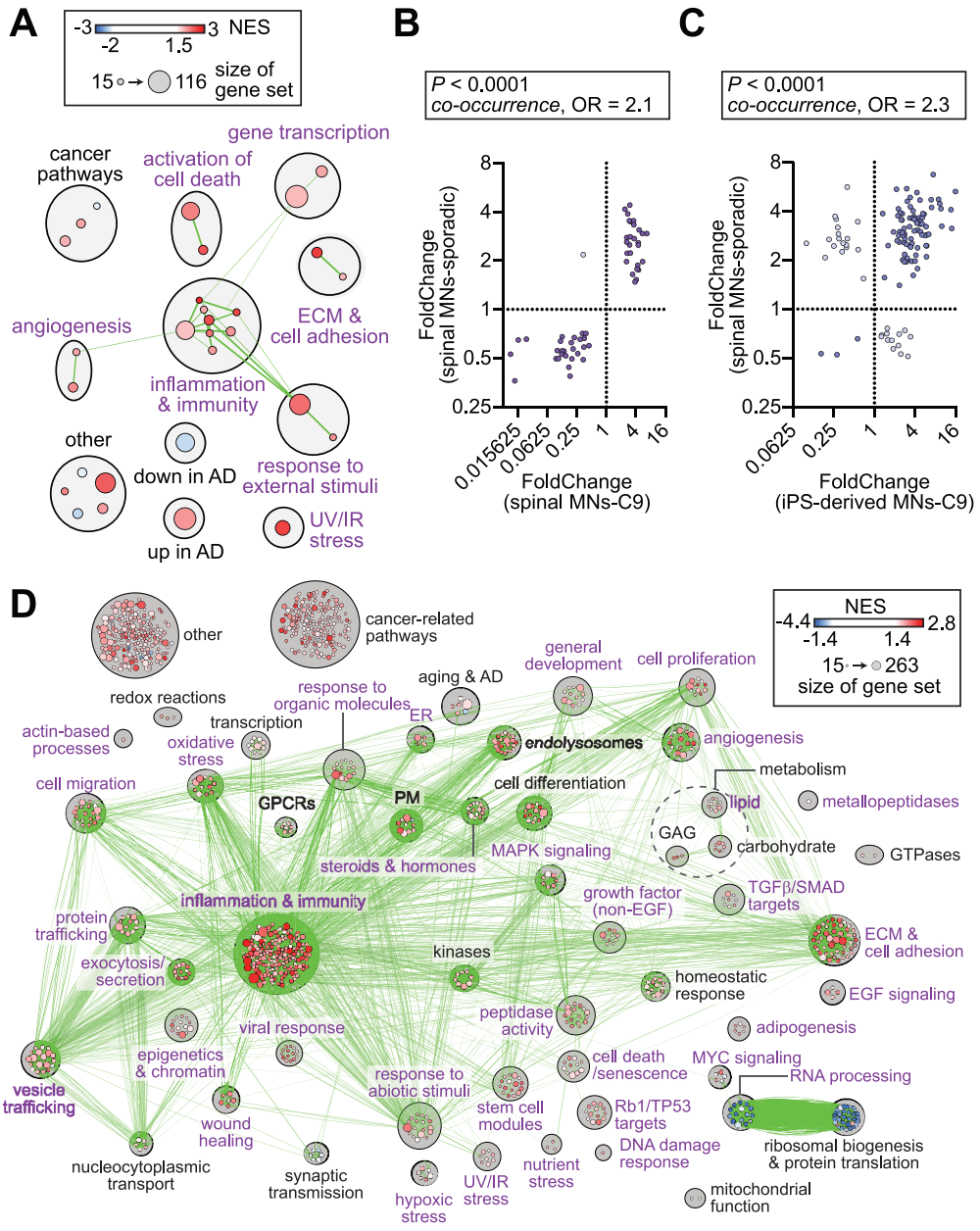

Figure 4. Analysis of differential gene expression in C9-ALS and sALS spinal motor neurons. (**A**) Pathway enrichment analysis of DEGs in C9-ALS spinal motor neurons. DEGs were ranked and subjected to GSEA. Gene sets with FDR $q < 0.1$ are shown as nodes. Node color represents NES from GSEA. Node size represents number of genes. Gene sets that are common between iPSC motor neurons and postmortem spinal motor neurons are labeled in *magenta*. (**B**) Scatterplot showing DEGs that co-occur in C9-ALS and sALS spinal motor neurons. Fold change of each DEG in C9-ALS was plotted against values for the same genes in sALS motor neurons. Contingency test results (Fisher's exact test; OR, odd ratio) are shown in the box. Co-occurring genes occupy the upper right and lower left quadrants and are shown in *magenta*. (**C**) Scatterplot showing co-occurring DEGs in sALS spinal motor neurons and C9-ALS iPSC motor neurons. Fold change of each DEG in C9-ALS iPSC-derived motor neurons was plotted against values for the same genes in sALS spinal motor neurons. Contingency test results (Fisher's exact test; OR, odd ratio) are shown in the box. Co-occurring genes occupy the upper right and lower left quadrants and are shown in *blue*. (**D**) Same as in (A) but with DEGs from sALS spinal motor neurons. Gene sets that are common between sALS and C9-ALS are labeled in *magenta*.

of downregulated gene sets. ECM and cell adhesion genes that were repressed in SOD1^{A272C}-ALS were upregulated in either C9-ALS or sALS and vice versa (Fig. 5B). In addition, modules pertinent to DREAM/E2F targets and microtubule-based transport—significantly repressed in C9-ALS patients (Fig. 2A)—were significantly upregulated in SOD1^{A272C}-ALS (Fig. 5A). In summary, transcriptional signatures associated with C9- and sALS are not only distinct from those in SOD1^{A272C}-ALS, some of the modules are diametrically opposite of those observed neurons bearing *SOD1* mutations. Concordantly, we found that DEGs in SOD1^{A272C}-ALS did not exhibit statistically significant co-occurrence in the C9-ALS or sALS data sets (Fig. 5C and D).

Next, we examined the transcriptional profile in iPSC-derived motor neurons generated from ALS patients that carry the *SOD1^{A4V}* mutation (10). Although a larger number of DEGs were upregulated in SOD1^{A4V}-ALS (Fig. 6A, inset), upregulated DEGs did not exhibit enrichment for gene sets (Fig. 6A). This was a clear difference from C9-ALS where a larger number of gene sets

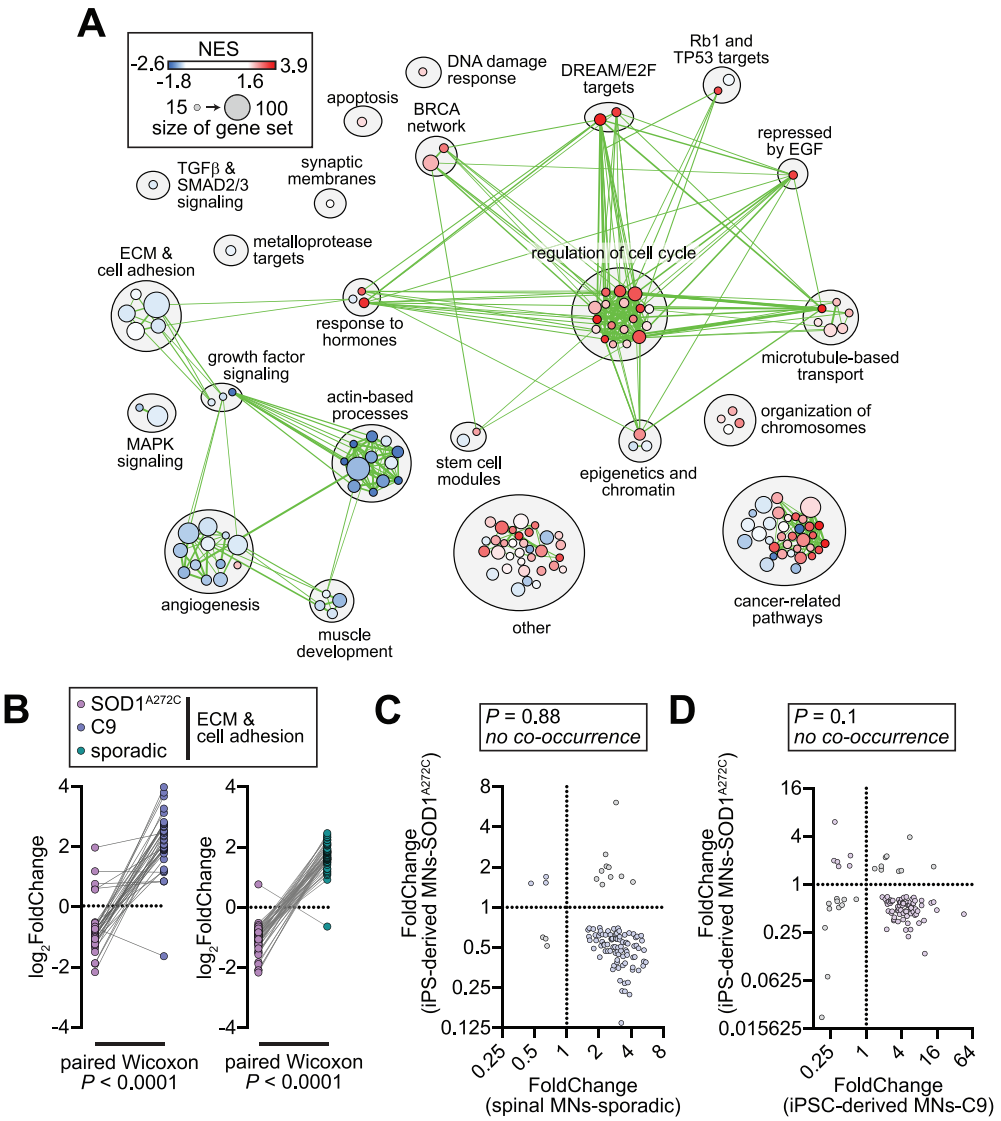

Figure 5. Distinct transcriptional signatures in iPSC motor neurons derived from SOD1^{A272C}-ALS patients. (**A**) Pathway enrichment analysis of DEGs in SOD1^{A272C}-ALS iPSC motor neurons. DEGs were ranked and subjected to GSEA. Gene sets with FDR $q < 0.1$ are shown as nodes. Node color represents NES from GSEA. Node size represents number of genes. (**B**) Comparison of relative expression levels of genes belonging to the ECM and cell adhesion module among SOD1^{A272C}-ALS and either C9-ALS iPSC-derived motor neurons or sALS spinal motor neurons. Expression levels of the same genes are connected by *grey* lines. Paired Wilcoxon tests were performed to examine differences in the distribution of the fold change values in the different data sets. (**C**) Scatterplot showing DEGs co-occurring in sALS spinal motor neurons and SOD1^{A272C}-ALS iPSC motor neurons. Fold change values of DEGs in sALS spinal motor neuron were plotted against values for the same genes in SOD1^{A272C}-ALS iPSC motor neurons. Contingency test results (Fisher's exact test) are shown in the box. (**D**) Scatterplot showing DEGs co-occurring in iPSC-derived motor neurons from C9-ALS and SOD1^{A272C}-ALS. Fold change values of DEGs in C9-ALS were plotted against values for the same genes in SOD1^{A272C}-ALS. Contingency test results (Fisher's exact test) are shown in the box.

had NES > 0. Furthermore, *SOD1*^{A4V}-expressing motor neurons did not exhibit an increase in the expression of genes that constitute modules related to ECM/cell adhesion or TGFβ/SMAD targets (Fig. 6A and Supplementary Material, Table S6). The only module that was upregulated in SOD1^{A4V}-ALS motor neurons were composed of gene sets related to cell cycle progression—a feature these neurons shared with SOD1^{A272C}-ALS motor neurons (Fig. 5A). Indeed, we detected significant co-occurrence of DEGs in the two SOD1 data sets (Fig. 6B).

DEGs in SOD1^{A4V}-ALS and C9-ALS showed statistically significant mutual exclusion (Fig. 6C), which underscores the robust differences between the transcriptomes in SOD1^{A4V}-ALS and C9-ALS. Genes that were upregulated in C9-ALS were repressed in SOD1^{A4V}-ALS and vice versa (Fig. 6C, inset). A module consisted of gene sets that are repressed during aging and in other neurodegenerative diseases was repressed in SOD1^{A4V}-ALS (Fig. 6A and Supplementary Material, Table S6). This observation suggests that SOD1^{A4V}-ALS shares ontological features with other neurodegenerative diseases. Furthermore, modules related to ribosomal biogenesis, RNA processing and mitochondrial function [mitochondrial translation, Tricarboxylic Acid Cycle (TCA) and Oxidative Phosphorylation (OXPHOS)] are similarly repressed in both SOD1^{A4V}-ALS and sALS (Figs 6A and 4D, respectively), which the observation that DEGs in SOD1^{A4V}-ALS and sALS exhibited significant co-occurrence (Fig. 6D). Therefore, sALS motor neurons exhibit transcriptional features that are observed in both C9-ALS and SOD1-ALS.

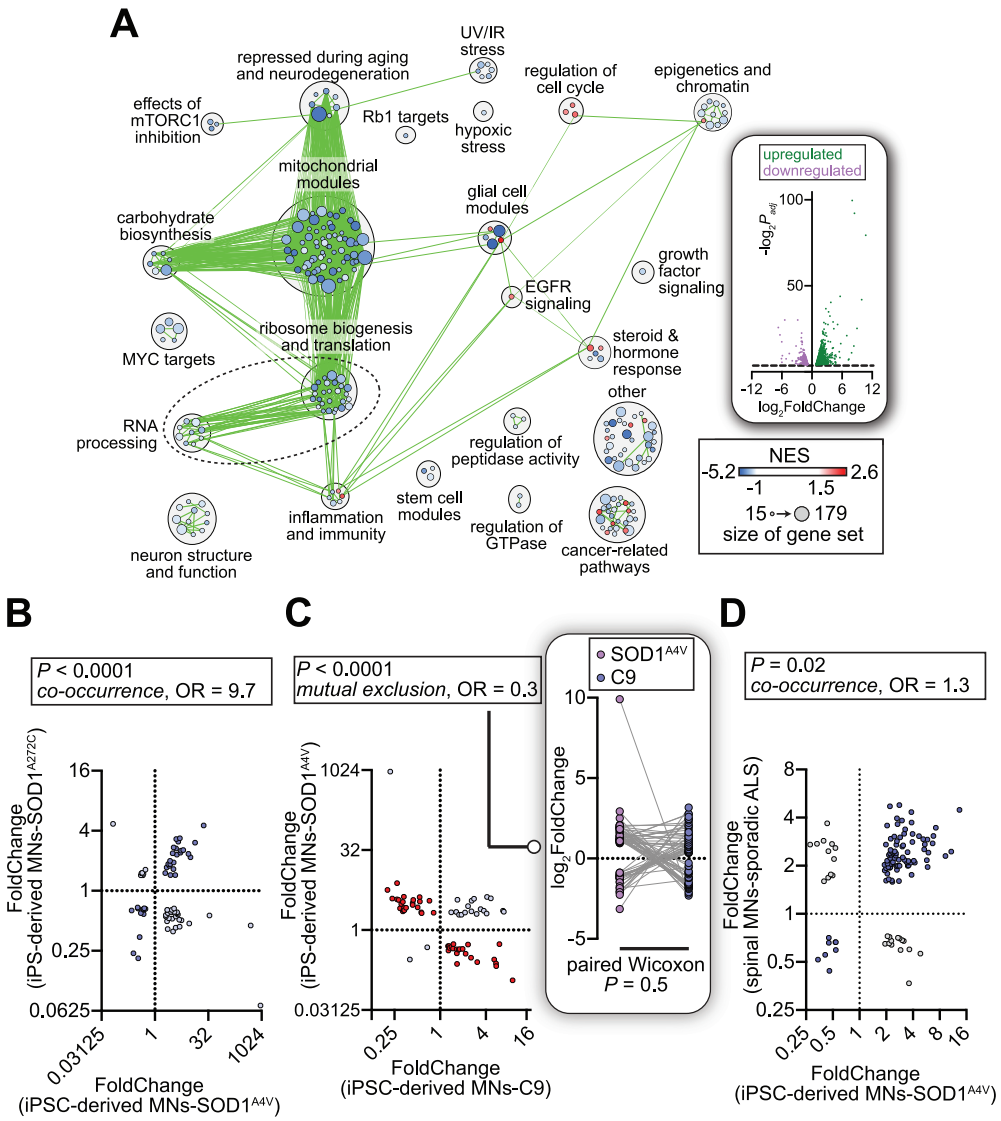

Figure 6. Distinct transcriptional signatures in iPSC motor neurons derived from SOD1^{A4V}-ALS patients. (**A**) Pathway enrichment analysis of DEGs in SOD1^{A4V}-ALS iPSC motor neurons. DEGs were ranked and subjected to GSEA. Gene sets with FDR $q < 0.1$ are shown as nodes. Node color represents NES from GSEA. Node size represents number of genes. Inset shows volcano plot of log values of the fold change in gene expression in SOD1^{A4V}-ALS motor neurons plotted against negative log values of the adjusted *P*-values. Horizontal *dashed line* indicates the cutoff for significance ($P_{adj} = 0.05$). Up- and downregulated genes are represented by *green* and *magenta* circles, respectively. (**B**) Scatterplot showing co-occurrence of DEGs in SOD1^{A4V}-ALS and SOD1^{A272C}-ALS iPSC-derived motor neurons. Fold change values of DEGs in SOD1^{A4V}-ALS were plotted against values for the same genes in SOD1^{A272C}-ALS. Contingency test results (Fisher's exact test; OR, odd ratio) are shown in the box. Co-occurring genes occupy the upper right and lower left quadrants and are shown in *blue*. (**C**) Scatterplot showing co-occurrence of DEGs in C9-ALS and SOD1^{A4V}-ALS iPSC-derived motor neurons. Fold change values of DEGs in C9-ALS were plotted against values for the same genes in SOD1^{A4V}-ALS iPSC motor neuron. Contingency test results (Fisher's exact test; OR, odd ratio) are shown in the box above the plot. Mutually exclusive genes occupy the upper left and lower right quadrants and are shown in *red*. Inset on the right has comparison of fold change in expression of the same genes in SOD1^{A4V}-ALS and C9-ALS iPSC-derived motor neurons. (**D**) Scatterplot showing co-occurrence of DEGs in SOD1^{A4V}-ALS and sALS. Fold change values of DEGs in SOD1^{A4V}-ALS iPSC-derived motor neurons were plotted against values for the same genes in sALS motor neurons. Contingency test results (Fisher's exact test; OR, odd ratio) are shown in box above the plot. Co-occurring genes occupy the upper right and lower left quadrants and are shown in *blue*.

Discussion

Independently generated data sets from C9-ALS and sALS motor neurons exhibit similar gene expression signatures

In this study, we undertook meta-analyses of transcriptomic and proteomic data sets to reveal alterations that occur in both C9-ALS and sALS. We detected increased expression of genes that encode ECM and cell adhesion proteins. We also identified elevated expression of other gene signatures such as regulators of angiogenesis and genes related to stress response and inflammation. Similarities in gene expression signatures in three independently generated data sets support the notion that alterations in the aforementioned pathways are significant to ALS pathology. Overlap in signatures between iPSC-derived and spinal motor neurons also indicates that transcriptional alterations in patient-derived neurons are not idiosyncratic features of the protocol used to generate those cells. Rather, our analyses lend credence to the idea that transcriptional signatures evident in iPSC-derived motor neurons mimic those occurring in the native environment in the spinal cord.

Modules of coregulated genes in C9-ALS and sALS motor neurons

Both C9-ALS and sALS data sets revealed differences in the expression of a modest number of previously identified ALS-related genes. For instance, iPSC-derived motor neurons showed elevated abundance of the mRNA and protein of the ALS-linked, *PRPH* (35–37) (Supplementary Material, Table S1A). This finding was intriguing given that prior studies have demonstrated age-dependent motor neuron degeneration and Tumor Necrosis Factor (TNF)α-dependent neuronal death following overexpression of *PRPH* (38–40). Spinal motor neurons from C9-ALS and iPSC-derived motor neurons from sALS exhibited decreased expression of *C9ORF92* and *SIGMAR1* and *VAPB*, respectively (Supplementary Materials, Tables S4A and S5). Diminished expression of these genes, especially *VAPB*, has been previously reported in sALS spinal cords (41).

iPSC-derived C9-ALS motor neurons exhibited diminished expression of DNA damage repair genes, which is consistent with the notion that ALS is associated with transcription-related DNA damage (42–45). PARP1—an important player in single-strand DNA repair via catalysis of poly (ADP-ribose) polymerization (46) and having potential roles in ALS (45,47–49)—was downregulated at both transcriptomic and proteomic levels in iPSC-derived motor neurons from C9-ALS patients (Supplementary Materials, Tables S1 and S4A). *PARP1* expression was also significantly diminished in spinal motor neurons from C9-ALS patients (Supplementary Material, Table S5). These data point to a consistent downregulation of *PARP1* in C9-ALS, which could indicate defects in DNA repair in the diseased motor neurons. iPSC-derived C9-ALS motor neurons also showed repression of genes related to microtubule-based transport, which could reflect defects in microtubule-dependent axonal trafficking, a common cellular phenotype in ALS (50).

GSEA indicated significant enrichment for genes whose expression is regulated by E2F4/DREAM complex or TGFβ/SMAD2/SMAD3 signaling. Whereas most E2F4/DREAM targets belonged to cluster 1, TGFβ/SMAD2/SMAD3 targets were enriched in cluster 5 (Supplementary Materials, Tables S1B and S2). Given the functions of E2F4/DREAM in gene repression and maintenance of quiescence (51–53), diminished transcription and protein levels of E2F4/DREAM targets suggest elevated transcription factor activity in sALS. In contrast, since SMAD2/SMAD3 are transcriptional activators, increased expression of target genes suggests enhanced TGFβ-signaling in ALS motor neurons. Interestingly, SMAD2/SMAD3 promote E2F4 association with target promoters (54,55). Thus, crosstalk between these two transcriptional modules is a possibility.

Signature of repressed mRNA processing in sALS

During ALS progression, RNA-binding proteins (RBPs) are sequestered by RNA granules resulting in a loss of RBP availability and an inexorable decline in neuronal function (16). Our analyses suggest that the cell biological changes may be 'imprinted' on the transcriptomic data. Alternatively, defective mRNA processing could arise from diminished expression of RBPs. If so, a decline in expression of genes related to mRNA processing could be an attempt to mitigate the aggregation of RNA granules. Because mRNA processing and DNA repair genes are under the control of common transcriptional regulators such as E2F4/DREAM (56), repression of the entire module to prevent the expression of RBPs bears the potential to collaterally suppress the DNA damage response leading to genomic instability. If so, the pathological consequences of module repression could be mitigated by targeting upstream regulators such as the E2F4/DREAM complex.

Closing remarks

At the level of outward phenotypic manifestations, many cases of ALS look alike. Given that different causal mutations often result in robust phenotypic similarities, it would be tempting to assume that targeting the cellular processes caused by any one mutation would be beneficial in all cases of ALS. However, our analyses suggest otherwise. While independently generated C9-ALS data sets show remarkably concordant transcriptional signatures, these signatures are conspicuously absent in cases of SOD1-ALS. The transcriptional signature of sALS, which may be caused by a plethora of unknown genetic lesions, lies between the extremes of C9- and SOD1-ALS. Other studies have also reported differences and similarities in the signatures in different types of ALS (5,6,16,31). Although additional studies will be needed to uncover the significance of the similarities and distinctions, it stands to reason that therapeutic approaches designed to mitigate alterations in one form of the disease could be ineffective or even worsen another form of the disease. Overall, our analyses inform the notion that targeting patient-specific alterations may be the best strategy when developing effective therapeutic strategies.

Materials and Methods

Transcriptomic data sets

RNA-Seq data sets generated from iPSC-derived motor neurons of C9-ALS and control patient cohorts were obtained from NeuroLINCS (http://neurolincs.org/data/; data set ID: LDS-1398). The cells were differentiated using the iMN long differentiation protocol. These data were generated by total RNA-Seq carried out in duplicate using RNA obtained from iPSC-derived motor neurons from four ALS and three control individuals (ALS samples—CS28iALS_iMNS, CS29iALS_iMNS, CS30iALS_iMNS, CS52iALS_iMNS; control samples—CS00iCTR_iMNS, CS25iCTR_iMNS, CS83iCTR_iMNS). Analyses of these patients and iPSC-derived motor neurons have been described elsewhere (11). The RNA-Seq protocol is available at the NeuroLINCS website.

Other RNA-Seq and microarray data sets were obtained from Gene Expression Omnibus (GEO). GEO accession numbers of the data sets were: laser-captured spinal motor neurons from postmortem C9-ALS (15)—GSE56504; laser-captured spinal motor neurons from postmortem sALS patient samples (32)—GSE103225; iPSC-derived motor neurons carrying $SOD1^{A272C}$ mutations (34)—GSE95089; and iPSC-derived motor neurons carrying $SOD1^{A4V}$ mutations (10)—GSE54409.

Proteomic data set

Proteomic data sets of iPSC-derived motor neurons from C9-ALS and controls were obtained from NeuroLINCS (data set ID: LDS-1424). These data were generated using the same samples that were used for RNA-Seq analyses and were thus composed of data from four ALS and three controls. Sum of the peptide counts of each protein ID protein was used as the measure of protein abundance. Protein IDs were mapped to their corresponding genes using UniProt ID mapping (https://www.uniprot.org/uploadlists/). A total of eight protein IDs (P12532, P32969, P62805, P68431, P84243, Q71DI3, Q96JG8 and Q9H3K6) were mapped to

multiple genes and were thus excluded from the analyses. Both P42166 and P42167 mapped to the same gene and were also excluded.

Analyses of differential gene expression and protein abundance

Differential gene expression analysis was performed using DESeq2 on the DEApp web-based application (https://yanli.shinyapps.io/DEApp) (19,20). For the transcriptomic data sets obtained from NeuroLINCS, we examined differential expression of only the protein-coding RNAs. First, we added the RNA-Seq from the technical replicates. The DEApp application removed the low expression terms with Counts Per Million (CPM) value ≤ 1 in less than two samples. $FDR < 0.1$ was used as the cutoff for identifying DEGs in the NeuroLINCS data set. In case of the proteomic data, we used geometric means of the technical replicates (57). DEPs were determined using LIMMA (58) with $FDR < 0.1$ as the cutoff, which provides high confidence results data sets with relatively few experimental replicates (59–61). For LIMMA, we used the DEBrowser web-based application (62).

Hierarchical clustering of DEGs

For each C9-ALS patient, we determined the relative change in expression of the 860 DEGs (*relative expression = expression in C9/average expression in controls*). Next, we determined the Pearson's correlation between each pair of DEGs. The resulting gene-by-gene matrix was composed of values ranging from -1 to 1. We converted these coefficient values to a 0 to 1 range using the formula: $C\_new = \left(C\_old + 1\right)/2$. Subsequently, we raised the values to the 6th power (equivalent to soft thresholding) to maximize the differences between separable clusters. We put the resulting adjacency matrix through an unsupervised hierarchical clustering algorithm (clusterMaker plugin of CytoScape) (63) to generate a heatmap. The parameters we used for clustering were clustering distance = correlation and clustering method = average.

GSEA

We performed GSEA using the standalone javaGSEA desktop application. Briefly, we first ranked a list of genes on the basis of differential gene expression in ALS. Next, we compared the ranked list either to a defined gene set or to an unbiased group of previously curated gene sets. Curated gene sets GSEA: C2 and GSEA: C5 were obtained from the Molecular Signatures Database (MSigDB). Data sets that met the $FDR < 0.1$ thresholds were considered significantly enriched.

Functional annotation of genes encoding differentially expressed proteins

To determine the functional ontologies associated with the proteins that are differentially expressed in C9-ALS motor neuron, we perform functional annotation of the mapped genes using DAVID (https://david.ncifcrf.gov/home.jsp) online application. DAVID provided information regarding enrichment of an ontological category over a background consisted of all human genes. We used only the categories with Benjamini $FDR < 0.05$ for analyses. We visualized the GSEA networks using the 'Enrichment map' plugin from CytoScape (64), with edge values being the similarity coefficients calculated by the plugin. Functional modules were annotated manually. The layout used for each module was, 'Edge-weighted spring-embedded layout' (65).

Statistical methods and data analyses

Graphs were generated using Prism (GraphPad). We determined significance using Student's *t*-tests. For data that were not normally distributed, we used Mann–Whitney test and Wilcoxon signed-rank test to compare medians to a hypothetical value. We compared odd ratios and used two-tailed Fisher's exact test to assess co-occurrence or mutual exclusivity. Corrections for multiple hypotheses testing were applied whenever appropriate.

Supplementary Material

Supplementary Material is available at *HMG* online.

Acknowledgements

We thank Drs. Jeffrey Chang and Marco Sardiello for useful discussions and Nicholas Karagas for critical comments on the manuscript.

Conflict of Interest statement. None declared.

Funding

National Institute of Neurological Disorders and Stroke (R01NS0 81301 and R21AG061646 to K.V.)

References

1. Taylor, J.P., Brown, R.H. and Cleveland, D.W. (2016) Decoding ALS: from genes to mechanism. *Nature*, **539**, 197–206.
2. Ling, S.-C., Polymenidou, M. and Cleveland, D.W. (2013) Converging mechanisms in ALS and FTD: disrupted RNA and protein homeostasis. *Neuron*, **79**, 416–438.
3. Renton, A.E., Majounie, E., Waite, A., Simón-Sánchez, J., Rollinson, S., Gibbs, J.R., Schymick, J.C., Laaksovirta, H., van Swieten, J.C., Myllykangas, L. *et al.* (2011) A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS–FTD. *Neuron*, **72**, 257–268.
4. DeJesus-Hernandez, M., Mackenzie, I.R., Boeve, B.F., Boxer, A.L., Baker, M., Rutherford, N.J., Nicholson, A.M., Finch, N.A., Flynn, H., Adamson, J. *et al.* (2011) Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron*, **72**, 245–256.
5. van Rheenen, W., van Blitterswijk, M., Huisman, M.H.B., Vlam, L., van Doormaal, P.T.C., Seelen, M., Medic, J., Dooijes, D., de Visser, M., van der Kooi, A.J. *et al.* (2012) Hexanucleotide repeat expansions in C9ORF72 in the spectrum of motor neuron diseases. *Neurology*, **79**, 878–882.
6. Tsai, C.-P., Soong, B.-W., Tu, P.-H., Lin, K.-P., Fuh, J.-L., Tsai, P.-C., Lu, Y.-C., Lee, I.-H. and Lee, Y.-C. (2012) A hexanucleotide repeat expansion in C9ORF72 causes familial and sporadic ALS in Taiwan. *Neurobiol. Aging*, **33**, 2232.e11–2232.e18.
7. Mok, K.Y., Koutsis, G., Schottlaender, L.V., Polke, J., Panas, M. and Houlden, H. (2012) High frequency of the expanded C9ORF72 hexanucleotide repeat in familial and sporadic Greek ALS patients. *Neurobiol. Aging*, **33**, 1851.e1–1851.e5.

8. Sabatelli, M., Conforti, F.L., Zollino, M., Mora, G., Monsurrò, M.R., Volanti, P., Marinou, K., Salvi, F., Corbo, M., Giannini, F. *et al.* (2012) C9ORF72 hexanucleotide repeat expansions in the Italian sporadic ALS population. *Neurobiol. Aging*, **33**, 1848.e15–1848.e20.
9. Egawa, N., Kitaoka, S., Tsukita, K., Naitoh, M., Takahashi, K., Yamamoto, T., Adachi, F., Kondo, T., Okita, K., Asaka, I. *et al.* (2012) Drug screening for ALS using patient-specific induced pluripotent stem cells. *Sci. Transl. Med.*, **4**, 145ra104–145ra104.
10. Kiskinis, E., Sandoe, J., Williams, L.A., Boulting, G.L., Moccia, R., Wainger, B.J., Han, S., Peng, T., Thams, S., Mikkilineni, S. *et al.* (2014) Pathways disrupted in human ALS motor neurons identified through genetic correction of mutant SOD1. *Cell Stem Cell*, **14**, 781–795.
11. Sareen, D., O'Rourke, J.G., Meera, P., Muhammad, A.K.M.G., Grant, S., Simpkinson, M., Bell, S., Carmona, S., Ornelas, L., Sahabian, A. *et al.* (2013) Targeting RNA foci in iPSC-derived motor neurons from ALS patients with a C9ORF72 repeat expansion. *Sci. Transl. Med.*, **5**, 208ra149–208ra149.
12. Bhinge, A., Namboori, S.C., Zhang, X., VanDongen, A.M.J. and Stanton, L.W. (2017) Genetic correction of SOD1 mutant iPSCs reveals ERK and JNK activated AP1 as a driver of neurodegeneration in amyotrophic lateral sclerosis. *Stem Cell Reports*, **8**, 856–869.
13. Rabin, S.J., Kim, J.M.'.H.'., Baughn, M., Libby, R.T., Kim, Y.J., Fan, Y., Libby, R.T., La Spada, A., Stone, B. and Ravits, J. (2010) Sporadic ALS has compartment-specific aberrant exon splicing and altered cell–matrix adhesion biology. *Hum. Mol. Genet.*, **19**, 313–328.
14. Kotni, M.K., Zhao, M. and Wei, D.-Q. (2016) Gene expression profiles and protein–protein interaction networks in amyotrophic lateral sclerosis patients with C9orf72 mutation. *Orphanet J. Rare Dis.*, **11**, 148.
15. Highley, J.R., Kirby, J., Jansweijer, J.A., Webb, P.S., Hewamadduma, C.A., Heath, P.R., Higginbottom, A., Raman, R., Ferraiuolo, L., Cooper-Knock, J. *et al.* (2014) Loss of nuclear TDP-43 in amyotrophic lateral sclerosis (ALS) causes altered expression of splicing machinery and widespread dysregulation of RNA splicing in motor neurones. *Neuropathol. Appl. Neurobiol.*, **40**, 670–685.
16. Conlon, E.G., Fagegaltier, D., Agius, P., Davis-Porada, J., Gregory, J., Hubbard, I., Kang, K., Kim, D., Phatnani, H., Kwan, J. *et al.* (2018) Unexpected similarities between C9ORF72 and sporadic forms of ALS/FTD suggest a common disease mechanism. *Elife*, **7**, e37754.
17. Bandyopadhyay, U., Cotney, J., Nagy, M., Oh, S., Leng, J., Mahajan, M., Mane, S., Fenton, W.A., Noonan, J.P. and Horwich, A.L. (2013) RNA-Seq profiling of spinal cord motor neurons from a presymptomatic SOD1 ALS mouse. *PLoS One*, **8**, e53575.
18. Gurney, M.E., Pu, H., Chiu, A.Y., Dal Canto, M.C., Polchow, C.Y., Alexander, D.D., Caliendo, J., Hentati, A., Kwon, Y.W. and Deng, H.X. (1994) Motor neuron degeneration in mice that express a human Cu,Zn superoxide dismutase mutation. *Science*, **264**, 1772–1775.
19. Li, Y. and Andrade, J. (2017) DEApp: an interactive web interface for differential expression analysis of next generation sequence data. *Source Code Biol. Med.*, **12**, 2.
20. Love, M.I., Huber, W. and Anders, S. (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.*, **15**, 550.
21. Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. *et al.* (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. USA*, **102**, 15545–15550.
22. Mootha, V.K., Lindgren, C.M., Eriksson, K.-F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E. *et al.* (2003) PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.*, **34**, 267–273.
23. Bae, B.-I., Xu, H., Igarashi, S., Fujimuro, M., Agrawal, N., Taya, Y., Hayward, S.D., Moran, T.H., Montell, C., Ross, C.A. *et al.* (2005) p53 mediates cellular dysfunction and behavioral abnormalities in Huntington's disease. *Neuron*, **47**, 29–41.
24. Lapresa, R., Agulla, J., Sánchez-Morán, I., Zamarreño, R., Prieto, E., Bolaños, J.P. and Almeida, A. (2019) Amyloid-ß promotes neurotoxicity by Cdk5-induced p53 stabilization. *Neuropharmacology*, **146**, 19–27.
25. Guo, X., Disatnik, M.-H., Monbureau, M., Shamloo, M., Mochly-Rosen, D. and Qi, X. (2013) Inhibition of mitochondrial fragmentation diminishes Huntington's disease–associated neurodegeneration. *J. Clin. Invest.*, **123**, 5371–5388.
26. Sung, M.-K., Porras-Yakushi, T.R., Reitsma, J.M., Huber, F.M., Sweredoski, M.J., Hoelz, A., Hess, S. and Deshaies, R.J. (2016) A conserved quality-control pathway that mediates degradation of unassembled ribosomal proteins. *Elife*, **5**, e19105.
27. Wyant, G.A., Abu-Remaileh, M., Frenkel, E.M., Laqtom, N.N., Dharamdasani, V., Lewis, C.A., Chan, S.H., Heinze, I., Ori, A. and Sabatini, D.M. (2018) NUFIP1 is a ribosome receptor for starvation-induced ribophagy. *Science (80-.)*, **360**, 751–758.
28. An, H. and Harper, J.W. (2018) Systematic analysis of ribophagy in human cells reveals bystander flux during selective autophagy. *Nat. Cell Biol.*, **20**, 135–143.
29. Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Res.*, **37**, 1–13.
30. Huang, D.W., Sherman, B.T. and Lempicki, R.A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.*, **4**, 44–57.
31. Prudencio, M., Belzil, V.V., Batra, R., Ross, C.A., Gendron, T.F., Pregent, L.J., Murray, M.E., Overstreet, K.K., Piazza-Johnston, A.E., Desaro, P. *et al.* (2015) Distinct brain transcriptome profiles in C9orf72-associated and sporadic ALS. *Nat. Neurosci.*, **18**, 1175–1182.
32. Krach, F., Batra, R., Wheeler, E.C., Vu, A.Q., Wang, R., Hutt, K., Rabin, S.J., Baughn, M.W., Libby, R.T., Diaz-Garcia, S. *et al.* (2018) Transcriptome–pathology correlation identifies interplay between TDP-43 and the expression of its kinase CK1E in sporadic ALS. *Acta Neuropathol.*, **136**, 405–423.
33. Conlon, E.G., Lu, L., Sharma, A., Yamazaki, T., Tang, T., Shneider, N.A. and Manley, J.L. (2016) The C9ORF72 GGGGCC expansion forms RNA G-quadruplex inclusions and sequesters hnRNP H to disrupt splicing in ALS brains. *Elife*, **5**, e17820.
34. Wang, L., Yi, F., Fu, L., Yang, J., Wang, S., Wang, Z., Suzuki, K., Sun, L., Xu, X., Yu, Y. *et al.* (2017) CRISPR/Cas9-mediated targeted gene correction in amyotrophic lateral sclerosis patient iPSCs. *Protein Cell*, **8**, 365–378.
35. Robertson, J., Doroudchi, M.M., Nguyen, M.D., Durham, H.D., Strong, M.J., Shaw, G., Julien, J.-P. and Mushynski, W.E. (2003) A neurotoxic peripherin splice variant in a mouse model of ALS. *J. Cell Biol.*, **160**, 939–949.
36. Gros-Louis, F., Larivière, R., Gowing, G., Laurent, S., Camu, W., Bouchard, J.-P., Meininger, V., Rouleau, G.A. and Julien, J.-P. (2004) A frameshift deletion in peripherin gene associated

with amyotrophic lateral sclerosis. *J. Biol. Chem.*, **279**, 45951–45956.
37. Leung, C.L., He, C.Z., Kaufmann, P., Chin, S.S., Naini, A., Liem, R.K.H., Mitsumoto, H. and Hays, A.P. (2004) A pathogenic peripherin gene mutation in a patient with amyotrophic lateral sclerosis. *Brain Pathol.*, **14**, 290–296.
38. Beaulieu, J.M., Nguyen, M.D. and Julien, J.P. (1999) Late onset of motor neurons in mice overexpressing wild-type peripherin. *J. Cell Biol.*, **147**, 531–544.
39. Robertson, J., Beaulieu, J.-M., Doroudchi, M.M., Durham, H.D., Julien, J.-P. and Mushynski, W.E. (2001) Apoptotic death of neurons exhibiting peripherin aggregates is mediated by the proinflammatory cytokine tumor necrosis factor-α. *J. Cell Biol.*, **155**, 217–226.
40. Xiao, S., Tjostheim, S., Sanelli, T., McLean, J.R., Horne, P., Fan, Y., Ravits, J., Strong, M.J. and Robertson, J. (2008) An aggregate-inducing peripherin isoform generated through intron retention is upregulated in amyotrophic lateral sclerosis and associated with disease pathology. *J. Neurosci.*, **28**, 1833–1840.
41. Anagnostou, G., Akbar, M.T., Paul, P., Angelinetta, C., Steiner, T.J. and de Belleroche, J. (2010) Vesicle associated membrane protein B (VAPB) is decreased in ALS spinal cord. *Neurobiol. Aging*, **31**, 969–985.
42. Hill, S.J., Mordes, D.A., Cameron, L.A., Neuberg, D.S., Landini, S., Eggan, K. and Livingston, D.M. (2016) Two familial ALS proteins function in prevention/repair of transcription-associated DNA damage. *Proc. Natl. Acad. Sci. USA*, **113**, E7701–E7709.
43. Sama, R.R.K., Ward, C.L. and Bosco, D.A. (2014) Functions of FUS/TLS from DNA repair to stress response: implications for ALS. *ASN Neuro*, **6**, 175909141454447.
44. Wang, W.-Y., Pan, L., Su, S.C., Quinn, E.J., Sasaki, M., Jimenez, J.C., Mackenzie, I.R.A., Huang, E.J. and Tsai, L.-H. (2013) Interaction of FUS and HDAC1 regulates DNA damage response and repair in neurons. *Nat. Neurosci.*, **16**, 1383–1391.
45. Mastrocola, A.S., Kim, S.H., Trinh, A.T., Rodenkirch, L.A. and Tibbetts, R.S. (2013) The RNA-binding protein fused in sarcoma (FUS) functions downstream of poly (ADP-ribose) polymerase (PARP) in response to DNA damage. *J. Biol. Chem.*, **288**, 24731–24741.
46. Benjamin, R.C. and Gill, D.M. (1980) ADP-ribosylation in mammalian cell ghosts. Dependence of poly (ADP-ribose) synthesis on strand breakage in DNA. *J. Biol. Chem.*, **255**, 10493–10501.
47. Rulten, S.L., Rotheray, A., Green, R.L., Grundy, G.J., Moore, D.A.Q., Gómez-Herreros, F., Hafezparast, M. and Caldecott, K.W. (2014) PARP-1 dependent recruitment of the amyotrophic lateral sclerosis-associated protein FUS/TLS to sites of oxidative DNA damage. *Nucleic Acids Res.*, **42**, 307–314.
48. Naumann, M., Pal, A., Goswami, A., Lojewski, X., Japtok, J., Vehlow, A., Naujock, M., Günther, R., Jin, M., Stanslowsky, N. *et al.* (2018) Impaired DNA damage response signaling by FUS-NLS mutations leads to neurodegeneration and FUS aggregate formation. *Nat. Commun.*, **9**, 335.
49. Wang, H., Guo, W., Mitra, J., Hegde, P.M., Vandoorne, T., Eckelmann, B.J., Mitra, S., Tomkinson, A.E., Van Den Bosch, L. and Hegde, M.L. (2018) Mutant FUS causes DNA ligation defects to inhibit oxidative damage repair in amyotrophic lateral sclerosis. *Nat. Commun.*, **9**, 3683.
50. De Vos, K.J. and Hafezparast, M. (2017) Neurobiology of axonal transport defects in motor neuron diseases: opportunities for translational research? *Neurobiol. Dis.*, **105**, 283–299.
51. Sadasivam, S. and DeCaprio, J.A. (2013) The DREAM complex: master coordinator of cell cycle-dependent gene expression. *Nat. Rev. Cancer*, **13**, 585–595.
52. Guarner, A., Morris, R., Korenjak, M., Boukhali, M., Zappia, M.P., Van Rechem, C., Whetstine, J.R., Ramaswamy, S., Zou, L., Frolov, M.V. *et al.* (2017) E2F/DP prevents cell-cycle progression in endocycling fat body cells by suppressing dATM expression. *Dev. Cell*, **43**, 689–703.e5.
53. Ren, B., Cam, H., Takahashi, Y., Volkert, T., Terragni, J., Young, R.A. and Dynlacht, B.D. (2002) E2F integrates cell cycle progression with DNA repair, replication, and G2/M checkpoints. *Genes Dev.*, **16**, 245–256.
54. Chen, C.-R., Kang, Y., Siegel, P.M. and Massagué, J. (2002) E2F4/5 and p107 as Smad cofactors linking the TGFβ receptor to c-myc repression. *Cell*, **110**, 19–32.
55. Yang, J., Song, K., Krebs, T.L., Jackson, M.W. and Danielpour, D. (2008) Rb/E2F4 and Smad2/3 link surviving to TGF-beta-induced apoptosis and tumor progression. *Oncogene*, **27**, 5326–5338.
56. Lee, B.-K., Bhinge, A.A. and Iyer, V.R. (2011) Wide-ranging functions of E2F4 in transcriptional activation and repression revealed by genome-wide analysis. *Nucleic Acids Res.*, **39**, 3558–3573.
57. Quackenbush, J. (2002) Microarray data normalization and transformation. *Nat. Genet.*, **32**, 496–501.
58. Smyth, G.K. (2004) Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Stat. Appl. Genet. Mol. Biol.*, **3**, 1–25.
59. Brusniak, M.-Y., Bodenmiller, B., Campbell, D., Cooke, K., Eddes, J., Garbutt, A., Lau, H., Letarte, S., Mueller, L.N., Sharma, V. *et al.* (2008) Corra: computational framework and tools for LC-MS discovery and targeted mass spectrometry-based proteomics. *BMC Bioinformatics*, **9**, 542.
60. Schwä, V., Rodríguez, I., León, L. and Jensen, O.N. (2013) Assessment and improvement of statistical tools for comparative proteomics analysis of sparse data sets with few experimental replicates. *Journal of Proteome Research*, **12**(9), 3874–3883. doi:10.1021/pr400045u.
61. Efstathiou, G., Antonakis, A.N., Pavlopoulos, G.A., Theodosiou, T., Divanach, P., Trudgian, D.C., Thomas, B., Papanikolaou, N., Aivaliotis, M., Acuto, O. *et al.* (2017) ProteoSign: an end-user online differential proteomics statistical analysis platform. *Nucleic Acids Res.*, **45**, W300–W306.
62. Kucukural, A., Yukselen, O., Ozata, D.M., Moore, M.J. and Garber, M. (2019) DEBrowser: interactive differential expression analysis and visualization tool for count data. *BMC Genomics*, **20**, 6.
63. Morris, J.H., Apeltsin, L., Newman, A.M., Baumbach, J., Wittkop, T., Su, G., Bader, G.D. and Ferrin, T.E. (2011) clusterMaker: a multi-algorithm clustering plugin for Cytoscape. *BMC Bioinformatics*, **12**, 436.
64. Merico, D., Isserlin, R., Stueker, O., Emili, A. and Bader, G.D. (2010) Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. *PLoS One*, **5**, e13984.
65. Kamada, T. and Kawai, S. (1989) An algorithm for drawing general undirected graphs. *Inf. Process. Lett.*, **31**, 7–15.