

Gene Expression Profiles in Sporadic ALS Fibroblasts Define Disease Subtypes and The Metabolic Effects of the Investigational Drug EH301

Jasmine A. Fels

Weill Cornell Graduate School of Medical Sciences

Gabriella Casalena

Weill Cornell Medical College: Weill Cornell Medicine

Csaba Konrad

Weill Cornell Medical College: Weill Cornell Medicine

Holly Holmes

Elysium Health

Ryan W. Dellinger

Elysium Health

Giovanni Manfredi (✉ gim2004@med.cornell.edu)

Joan and Sanford I Weill Medical College of Cornell University <https://orcid.org/0000-0003-3893-1348>

Research article

Keywords: ALS, fibroblasts, EH301, nicotinamide riboside, pterostilbene, weighted gene co-expression network analysis

Posted Date: October 13th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-952505/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Majority of ALS cases are sporadic (sALS), as they lack defined genetic causes. Metabolic alterations shared between the nervous system and skin fibroblasts have emerged in ALS. Recently, we found that a subgroup of sALS fibroblasts (sALS1) is characterized by metabolic profiles (metabotype) distinct from other sALS cases (sALS2) and controls, suggesting that metabolic therapies could be effective in sALS. The metabolic modulators nicotinamide riboside and pterostilbene (EH301) are under clinical development for the treatment of ALS. Here, we studied the metabolome and transcriptome of sALS cells to understand the molecular bases of sALS metabotypes and the impact of EH301.

Methods: Six fibroblast cell lines (3 male and 3 female subjects of similar ages) were used for each group (sALS1, sALS2, and controls). Metabolomics and transcriptomics were investigated at baseline and after EH301 treatment. Differential gene expression (DEGs) and metabolite abundance were assessed by a Wald Test and ANOVA, respectively, with FDR correction, and pathway analyses were performed. EH301 protection against metabolic stress was tested by thiol depletion. Weighted gene co-expression network analysis (WGCNA) was used to investigate the association of metabolic and clinical features and was also performed on the Answer ALS dataset from induced motor neurons (iMN). A machine learning model based on DEGs was tested as a sALS disease progression predictor.

Results: We found that the sALS1 transcriptome is distinct from sALS2 and that EH301 modifies gene expression differently in sALS1, sALS2, and controls. Furthermore, EH301 had strong protective effects against metabolic stress, which is linked to anti-inflammatory and antioxidant pathways. WGCNA revealed that ALS functional rating scale and metabotypes are associated with gene modules enriched for cell cycle, immunity, autophagy, and metabolism terms, which are modified by EH301. Meta-analysis of publicly available transcriptomics data from iMNs confirmed functional associations of genes correlated with disease traits. A small subset of genes differentially expressed in sALS fibroblasts could be used in a machine learning model to predict disease progression.

Conclusions: Multi-omics analyses of patient-derived fibroblasts highlighted differential metabolic and transcriptomic profiles in sALS metabotypes, which translate into differential responses to the investigational drug EH301.

Background

ALS (amyotrophic lateral sclerosis) is a rapidly progressive neurodegenerative disease involving upper and lower motor neurons. While only 10%-15% of the total cases are associated with known genetic mutations, most ALS patients have no familial history and no clear genetic alterations (sporadic ALS, sALS), [1]. While familial ALS-associated mutations are found in genes involved in several key cellular mechanisms, such as RNA binding, proteostasis, mitochondrial function, and cytoskeletal organization [2], the pathological mechanisms leading to sALS are still largely unknown. Therefore, there is a lack of sALS biomarkers and specific therapeutic targets. Although the clinical course in sporadic and familial cases reaches a generally predictable outcome, variability between patients has been recognized as an important factor in ALS, whose onset and progression are probably the result of complex interactions between genome, epigenome, and environment. Moreover, it is plausible that the disease is triggered by a variety of molecular abnormalities, ultimately converging onto common pathogenic pathways that lead to motor neuron death. Thus, the complexity and heterogeneity of sALS pathogenesis are probable causes for the failure of numerous clinical trials, so that currently there are only two approved drugs, Riluzole and Edavarone, which have modest clinical effects. As sALS heterogeneity has become evident, precision medicine approaches are gaining increasing attention. The ability to stratify ALS patients could inform the design of more precisely targeted therapeutic approaches and increase the probability of finding effective treatments for specific groups of patients. Understanding the differences and similarities among sALS patients at the molecular level could also contribute to the discovery of biomarkers that would improve reliability of trial endpoints and potentially provide a foundation for patient stratification and early intervention.

A recurring observation in ALS patients is the dysregulation of energy metabolism, with increased energy consumption and loss of fat mass, which sometimes even precedes disease onset [3]. In addition, increased glucose metabolism was described in the central nervous system of ALS patients [4]. Importantly, hypermetabolism in ALS is associated with greater functional decline and accelerated mortality [5]. Indeed, there is mounting evidence that metabolic alterations in ALS patients are shared between affected cells of the nervous system and other cell types. Specifically, we reported aberrantly increased energy metabolism in sALS patient-derived primary skin fibroblasts [6, 7]. Furthermore, our recent work showed that a subset of sALS fibroblasts is characterized by a distinct metabolic profile, that we defined as “sALS1” metabotype, which is associated with accelerated transsulfuration pathway for glutathione synthesis and glucose hypermetabolism [8]. The sALS1 metabotype was also identified based on targeted metabolomics in human plasma, indicating a direct relationship between skin fibroblasts and systemic metabolism [8]. These observations suggest that sALS patients could be stratified based on metabotypes, which may correspond to different pathogenic mechanisms and susceptibility to therapeutic interventions targeting specific metabolic alterations.

Energy metabolism alterations are often accompanied by oxidative stress and modification of the redox state of the cell, which is largely controlled by the levels of nicotinamide adenine dinucleotide (NAD). NAD functions as an electron carrier in many redox reactions and as a co-substrate for poly (ADP-ribose) polymerase (PARPs), ADP-ribosyl cyclases, and sirtuins. Therefore, NAD is a crucial element in the coordinated signaling between DNA, proteins, and metabolism [9]. NAD declines with age [10] and there is evidence of NAD metabolism imbalance in ALS patients and animal models [11]. Boosting NAD availability is considered a viable approach to restore or sustain cell metabolism, and it has been shown to be protective in *in vitro* [12] and *in vivo* ALS models [11]. In particular, the NAD precursor nicotinamide riboside (NR) has been deemed safe and effective in increasing circulating NAD levels in humans [13].

In addition to boosting NAD levels, another approach to modulate metabolism is through administration of polyphenols, a class of molecules with known antioxidant properties and protective effects against diseases associated with aging [14]. Pterostilbene (PT) is a polyphenol analog of resveratrol, but with improved bioavailability [15], which was shown to be protective in animal models of metabolic diseases [16] and neurodegeneration [17].

Recently, it was shown that the combination of NR and PT increases survival and delays motor neuron degeneration in the SOD1G93A mouse model of familial ALS [18]. Further, the therapeutic potential of a NR and PT combination (EH301) was suggested by a pilot clinical study, showing slower decline and improved muscle function in a small cohort of ALS patients [19]. Based on these promising early results, a larger two-dose, randomized, double-blind one-year trial has been initiated (The NO-ALS Study: A Trial of Nicotinamide Riboside/Pterostilbene Supplement in ALS [20]).

In this study, to better understand the molecular bases of sALS metabotypes and the impact of EH301 on sALS metabolism, we analyzed the metabolome and transcriptome of fibroblasts from patients with different sALS metabotypes and control individuals, before and after treatment with EH301. Furthermore, we performed weighted gene co-expression network analysis (WGCNA) in transcriptomic data from sALS fibroblasts. These analyses identified specific gene modules that correlated with clinical features and were modified by EH301 treatment. To confirm the fibroblasts results in cell types affected by the disease, we performed WGCNA in a publicly available transcriptomic dataset from induced motor neurons (iMNs). Lastly, we performed a proof-of-concept experiment to test the ability of a machine learning model to predict disease progression based on the expression of a few genes differentially expressed in sALS fibroblasts.

Methods

Cell culture

A total of eighteen primary fibroblasts lines derived from healthy donors or ALS patients (Table 1) were maintained in culture as previously reported [7]. sALS subclasses (sALS1 and sALS2) were defined based on previously published

metabolic profiles [8]. For experiments, cells were plated in Dulbecco's modified Eagle's medium (DMEM) medium containing 5 mM glucose and 2 mM glutamine, 10% FBS, 1% of 100× antibiotic/anti-mycotic (sterile filtered 10,000 units penicillin, 10 mg streptomycin and 25 µg amphotericin B per mL, and 2.5 µg/ml Plasmocin). Methionine and cystine depleted DMEM contained 10 mM glucose and 2 mM glutamine, 10% FBS, 1% of 100× antibiotic/anti-mycotic (sterile filtered 10,000 units penicillin, 10 mg streptomycin and 25 µg amphotericin B per mL, and 2.5 µg/ml Plasmocin). Cells were assessed at comparable passage number. EH301 was defined as the combination of 1 mM nicotinamide riboside (NR) and 10µM pterostilbene (PT) for the time specified in the text and figure legends.

Table 1
Characteristics of sALS and control subjects at time of biopsy.

	ID	Age at biopsy	Sex	Disease duration at time of skin biopsy (months)	ALSFRS-R total at biopsy	Rate of ALSFRS-R decline*	FVC %
Control	1	62	F				
	2	63	M				
	3	53	F				
	4	79	M				
	5	60	F				
	6	73	M				
sALS1	7	67	F	24	34	0.58	97
	8	73	M	10	24	2.4	38
	9	67	F	14	35	0.93	78
	10	64	M	24	38	0.42	55
	11	66	F	5	40	1.6	118
	12	56	M	34	39	0.26	81
sALS2	13	72	M	12	39	0.75	105
	14	73	F	6	39	1.5	59
	15	66	F	11	43	0.45	102
	16	68	M	32	32	0.5	78
	17	69	F	15	29	1.27	96
	18	55	M	10	33	1.5	116
* % Calculated as ((48 - ALSFRS at Skin BX) / Disease Duration at Skin BX)							

RNAseq and Quantitative Expression Analysis

500ng of total RNA extracted with TRIzol (Invitrogen) were used by the Weill Cornell Genomics Resources Core Facility to prepare 3'RNAseq libraries using the Lexogen QuantSeq 3' mRNA-Seq Library Prep Kit FWD for Illumina. The libraries were quantified on a Molecular Devices Spectra Max M2 plate reader (with the intercalating dye QuantiFluor) and pooled accordingly for maximum evenness. The pool was quantified by digital PCR and sequenced on 1 lane of an Illumina NextSeq500 sequencer, single end 1x86bp, and de-multiplexed based upon six base i7 indices using Illumina bcl2fastq2 software (version 2.18; Illumina, Inc., San Diego, CA).

Illumina adapters were removed from the de-multiplexed fastq files using Trimmomatic version 0.36 [21]. The trimmed reads were aligned to the human genome assembly GRCh38.p13 using the STAR aligner version 2.7.0f [22]. The output SAM files were converted to BAM using SAMtools version 1.8 [23], and the number of reads overlapping each gene in the gff3 file on the forward strand were counted using HTSeq-count version 0.6.1 [24]. The R package DESeq2 version 1.24.0 [25] was used to obtain both normalized and variance stabilized counts and to identify genes that were differentially expressed between controls and sALS lines. Genes with less than 10 total raw reads were filtered out before running the DESeq2 model, and all other filtering parameters were kept as DESeq2's defaults. Pathway analysis was performed with the free web tool WebGestalt [26] and the gprofiler2 package [27], and the cutoff for significance was a FDR corrected p-value < 0.05.

Metabolomics

Targeted metabolomics was performed by the Proteomics and Metabolomics Core Facility at Weill Cornell Medicine in New York. Metabolites were rapidly extracted in 80% ice-cold methanol; samples were cleared by centrifugation at 14,000 x g for 20 min at 4°C and stored at -80°C until analysis. Total protein was used for normalization, and relative abundance data were normalized with a log transformation and analyzed with the free online tool MetaboAnalyst 5.0 [28].

Cell viability

To measure viability, cells were plated in methionine and cystine depleted medium on 96 well glass bottom plates (Cellvis). After 72h, cells were stained with HOECHST 33342 (Invitrogen), Calcein AM (Invitrogen), Propidium Iodide (Invitrogen) for 30 minutes at 37°C. Fluorescence was acquired with the ImageXpress Pico Automated Imaging System. Viability threshold was determined by a combination of calcein and propidium iodide signals as previously reported [29].

Weighted gene co-expression network analysis (WGCNA)

WGCNA has been previously described [30, 31]. In brief, we used normalized gene expression data from DESeq2 as input for the functions included in the WGCNA package available from CRAN. iMN data was additionally normalized to remove batch effects using the `limma::removeBatchEffect()` function [32]. We generated a topological overlap matrix (TOM) based on expression values from 17,662 genes in the fibroblast dataset and 22,663 genes in the iMN dataset and identified modules using a dynamic tree cutting algorithm based on hierarchical clustering of TOM dissimilarity values. After optimization to maintain scale-free topology, we set the ideal soft power threshold value at 4 for the vehicle network, 20 for the EH301 network, and 8 for the iMN network. To allow for direct comparisons between the vehicle and EH301 networks, Q-Q scaling was performed such that the 95% quantiles of both matrices matched. For all networks the module merging parameter was kept consistent at 80%. Pairwise Pearson's correlations were used to calculate associations between modules and disease traits. Pathway analysis was done in the same way as described for RNAseq. Hierarchical clustering of ALSFRS-R and gene expression was done using the `dendextend` package [33]. Correlation between dendograms was calculated using Baker's gamma index, and significance was determined using a permutation test [34]. Logistic regression using the `glm()` function was used to calculate a model predicting ALSFRS-R progression. To evaluate model performance, we used a ten-fold cross validation approach, in which we randomly divided samples into ten sets. In each iteration, one set was used as the test data and the model was trained on the remaining 9 sets. Model performance was evaluated by calculating precision, accuracy, false positive rate, and false negative rate with the `caret` package [35], and with receiver operating characteristic analysis done with the `ROCR` package [36].

Statistical analyses

A Wald test was used to determine statistical significance of differential gene expression, with the cutoff being a False Discovery Rate of < 5% after Benjamini-Hochberg correction. Significance for differential metabolite abundance was determined with one-way ANOVA with post-hoc t-tests, with the cutoff being a False Discovery Rate of < 5% after Benjamini-Hochberg correction. All data visualization was done in R using the ggplot2, pheatmap, corrplot, and vennDiagram packages available from CRAN. Z scores were calculated from normalized counts for each gene using the standard formula $(x-\mu)/\sigma$, where x is the sample value, μ is the population mean, and σ is the population standard deviation.

Results

Transcriptomic analysis reveals higher number of DEGs in sALS1 than sALS2 fibroblasts and different transcriptional responses to EH301

To characterize gene expression profiles of sALS subgroups we performed 3'RNAseq analysis on fibroblast lines from control, sALS1, and sALS2 subjects (n=6 per group). In sALS1, 281 genes were differentially expressed (DEGs, $p_{adj} < 0.05$) relative to controls (Supplementary Table 1), while only one gene reached statistical significance in the comparison between sALS2 and controls (Fig. 1A, Supplementary Table 2). Interestingly, several genes relevant to neuronal function and development were differentially expressed in sALS1. For example, stathmin 2 (STMN2) was downregulated in sALS1 by ~80%. STMN2 has been linked to TDP-43 dysfunction [37] and a novel STMN2 genetic variant has been associated with ALS risk, onset, and progression [38]. Furthermore, the most upregulated gene in sALS1 was an antisense RNA for kinesin family member 5C (KIF5C-AS1). KIF5C is highly expressed in the brain and enriched in motor neurons [39], where it regulates axonal transport [40], and alterations of KIF5C are associated with intellectual disabilities and cortical development malformations [41, 42]. In addition, YIF1A, which was upregulated in sALS1, interacts with the ALS8 related protein VAPB [43] involved in neuronal ER-Golgi interactions [44]. HIST1H4C, a replication-dependent component of the nucleosome, was among the top downregulated genes in sALS1. Mutations affecting lysine 91 in HIST1H4C have been associated with a syndrome characterized by developmental anomalies and intellectual disabilities, indicating the importance of chromatin organization for the correct development and function of the nervous system [45]. SOX9, a transcription factor that controls several aspects of neurodevelopment [46] and is highly expressed in astrocytes and neural progenitor cells [47], was downregulated in sALS1. SYNE2 (nesprin), involved in organellar subcellular organization [48] and associated with muscular dystrophy [49], was also downregulated in sALS1. The only gene significantly upregulated in sALS2 compared to controls was MRE11, which encodes a double-strand break repair protein implicated in DNA damage response [50].

In addition to examining individual genes expression, we performed pathway analysis by Webgestalt [26] of biological processes and molecular function of sALS1 DEGs, which revealed that upregulated genes are involved in vesicular and protein transport and in extracellular matrix organization (Fig. 1B, Supplementary Table 3). Among the downregulated genes, the most enriched pathways in sALS1 were linked to cell cycle progression and cytoskeletal function (Fig. 1B).

Next, we evaluated the effects of EH301 on fibroblast transcriptomic profiles. Cells were exposed to EH301 (NR 1mM, PT 10 μ M) for 48 hours prior to RNA extraction. Interestingly, we found that EH301 affected a larger number of genes in sALS1 and sALS2 fibroblasts compared to controls (233 genes in sALS1, 202 genes in sALS2, 77 genes in controls), with little overlap between groups (Fig. 1C, Supplementary Tables 4-6). In control fibroblasts, pathway analysis of DEGs between vehicle and EH301 treated cells indicated that EH301 modified the expression of genes involved in mRNA processing (Fig. 1D, Supplementary Tables 7-9). For example, SYNCRIP, a ribonucleoprotein involved in RNA stabilization and editing, which has been associated with intellectual disabilities [51–53], was downregulated by EH301. CWC22 and CWC27, which cooperate during spliceosome assembly and are linked to developmental defects [54], were also downregulated in control fibroblasts by EH301. In sALS1 fibroblasts, EH301 influenced downstream steps of protein biosynthesis, by modifying the expression of genes involved in ribosome organization, translation initiation, and protein localization (Fig. 1D). Several

genes encoding ribosomal proteins, components of the 60S and the 40S subunits, were upregulated. Furthermore, two elements of the eukaryotic initiation factor 3 complex were differentially expressed after treatment. EIF3F, a positive regulator of NOTCH signaling [55], was upregulated. Conversely, EIF3J, involved in the recognition of starting codons [56] and in ribosome recycling [57], was downregulated. SEC11A, which mediates import of nascent protein into the ER [58], was upregulated, while KDELR3, mediating protein trafficking from Golgi to ER and involved in stress response [59], was downregulated by EH301. BCAP31, a chaperone abundant in the ER and involved in transmembrane protein export [60, 61] and in the assembly of mitochondrial Complex I [62], was upregulated. Both KDELR3 and BCAP31 have been associated with pathologies of the nervous system [63, 64]. Of note, treatment with EH301 normalized YIF1A expression in sALS1, while STMN2 and KIF5C-AS1 expression remained altered. Surprisingly, no pathway was found to be significantly enriched in sALS2 fibroblasts, even though the expression of 187 genes was altered by EH301 in this group.

In summary, RNAseq in human primary fibroblasts confirmed that, based on the number of DEGs, sALS1 samples are more distinct from controls than sALS2 and that genes involved in neurodevelopment and neuronal function are differentially regulated in sALS1 fibroblasts. Moreover, pathway analysis indicates that EH301 affects sALS1, sALS2, and control fibroblast gene expression differently, mostly affecting mRNA splicing and stability in controls and protein biosynthesis and localization in sALS1, while no specific pathways were identified in EH301 treated sALS2.

The metabolite profiles of sALS and control fibroblasts are modified by EH301

Next, we investigated how treatment with EH301 affects the metabolite profiles of sALS1, sALS2, and control fibroblasts. We performed targeted metabolomics in the same cell lines (n=6 per group) used for transcriptomics, in the same cell culture conditions. Following exclusion of low abundance hits, 166 metabolites were used for analysis. Metabolomics profiles showed that sALS1 had reduced cystathionine and increased betaine compared to controls (Supplementary Tables 10-12), corroborating previously reported differences in the transsulfuration pathway [8]. Cystathionine levels were unchanged in EH301 treated sALS1 but decreased in both control and sALS2 fibroblasts (Fig. 2A). Oxidized glutathione was increased after treatment in controls, but unchanged in sALS1 and sALS2 fibroblasts (Fig. 2A). Together, these results indicate that EH301 modulates the transsulfuration pathway, but does not correct the alterations observed in sALS1. As expected, metabolic pathway analysis showed that EH301 modifies metabolites of the nicotinate and nicotinamide pathway in all groups, increasing availability of NAD precursors and NAD (Fig. 2A, B). EH301 treated sALS2 also showed decreased fumarate and malate compared to vehicle treated cells, suggesting that NAD derived from NR accelerates the TCA cycle (Fig. 2A). D-glyceraldehyde-3-phosphate was significantly increased at baseline in sALS2 compared to controls and was normalized by EH301 (Fig. 2A), pointing to accelerated flux of NAD-dependent reactions in EH301 treated sALS2 cells. Riboflavin, the precursor of flavin mononucleotide (FMN) and flavin adenine dinucleotide (FAD), was decreased at baseline in both sALS1 and sALS2 fibroblasts and returned to control levels after treatment with EH301 (Fig. 2A). The pentose phosphate pathway was also affected by EH301 in control, sALS1, and sALS2 fibroblasts, with increased concentration of ribose and phosphorylated ribose in treated cells (Fig. 2A, B). Pyrimidine metabolism was modified by EH301 only in sALS1 and sALS2 fibroblasts, while controls were not affected (Fig. 2A, B). Purine metabolism, which was reported to be altered in sALS1 at baseline [8], was affected by EH301 in control and sALS1 cells (Fig. 2A, B). Interestingly, EH301 increased 1-methyladenosine in all groups. 1-methyladenosine is a S-adenosylmethionine (SAM)-dependent modification of RNA regulating mRNA localization, stability, translation, and splicing [65]. 1-Methyladenosine also responds to stress, decreasing upon glucose or amino acid starvation and increasing after heat shock [65]. 1-Methyladenosine can also modify tRNA, regulating its stability and folding [66] and is found in mtDNA-encoded transcripts [67]. Thus, some of EH301 effects on transcription and metabolism could be mediated by regulation of SAM-dependent epigenetics marks on RNA.

The alanine, aspartate, and glutamate metabolic pathways were globally altered by EH301 in sALS1 and sALS2 fibroblasts (Fig. 2A, B). Specifically, glutamate was decreased at baseline in sALS1, but not in sALS2, compared to controls (Fig. 2A, B). This selective glutamate decrease in sALS1 could be due to increased extrusion of glutamate in exchange for cystine

by the SLC7A11 transporter, which was shown to be upregulated in sALS1 fibroblasts [8]. EH301 decreased glutamate levels in all groups, a potentially disease-relevant effect, since extracellular glutamate homeostasis is known to be dysregulated in ALS [68].

EH301 strongly protects fibroblasts from cell death induced by thiol group depletion

Although EH301 did not correct the characteristic unbalances of transsulfuration metabolites in sALS1, we investigated if increased NAD availability and modification of amino acids metabolism by EH301 could improve cell viability under metabolic stress induced by methionine and cystine deprivation. To this end, we cultured cells for 72h in methionine and cystine depleted medium, in the presence or absence of EH301 or its individual components (NR or PT). Depletion of methionine and cystine caused similar levels of cell death in all groups (Fig. 2C, D). Interestingly, addition of NR alone was effective in protecting sALS1 fibroblasts from cell death, while the viability of sALS2 and controls was not improved (Fig. 2C, D). Treatment with PT alone was sufficient to prevent cell death in all groups (Fig. 2C, D). The combination of NR and PT (EH301) had comparable effects to PT alone. These results further indicate that the metabolic alterations of sALS1 are different than sALS2 and potentially more responsive to nicotinamide derivatives. They also indicate that PT is the most potent compound in protecting cells from the profound redox stress deriving from thiol group depletion.

Transcriptomic analysis performed on control cells exposed to medium depleted of methionine and cystine showed that treatment with EH301 reduces expression of genes involved in inflammation and apoptosis (Fig. 2E, F, Supplementary Tables 13-14). Of note, the expression of the stress response factor ATF3 was downregulated by EH301, while the levels of ATF5, which promotes expression of chaperones and pro-survival factors [69], was increased (Fig. 2E). Members of the kinesin family and TP53 were among the genes upregulated by EH301 in fibroblasts grown in absence of methionine and cystine. PMRT1 and PRMT2, regulating DNA damage response and other signaling pathways through SAM-dependent arginine methylation, were also increased by EH301. On the other hand, interleukins (CXCL1, CXCL2, CXCL3, CXCL5, CXCL6, CXCL8) were downregulated by the treatment. The transcripts of SOD2 and different metallothionein isoforms were reduced in fibroblasts treated with EH301 compared to vehicle, further indicating that EH301 acts through antioxidant and anti-inflammatory mechanisms, which prevent the need for upregulation of free radical scavengers and stress response genes under thiol-depleted conditions.

Weighted gene co-expression network analysis highlights associations between fibroblast transcriptional profiles and ALS clinical traits, which are altered by EH301 treatment

Weighted gene co-expression network analysis (WGCNA) is a powerful unbiased method for analysis of transcriptome-wide changes due to disease state [30, 70]. WGCNA differs from more traditional differential gene expression analysis methods, because it considers groups of genes with highly similar expression patterns across samples as part of a set of interconnected modules, rather than considering genes as single entities. This type of analysis increases the statistical power available to identify significant associations with phenotypic traits by minimizing noise. It may also provide more comprehensive information on complex biological processes [71].

To apply the WGCNA framework to our fibroblast gene expression dataset, we first constructed a co-expression matrix using normalized expression data for 17,662 genes. For this analysis we included all the vehicle-treated lines except one sALS1 that was identified as an outlier, based on its extreme distance from all other samples in hierarchical clustering (Fig. S1). We applied the same method to construct a matrix using all the EH301-treated samples. The WGCNA framework uses this matrix as input for hierarchical clustering to group highly co-expressed genes into modules. We identified 25 such modules in the vehicle network and 90 in the EH301 network (Fig. 3A, B). We next correlated module gene expression with 6 disease traits (disease status - ALS or control -, disease subgroup - sALS1 or sALS2 -, disease duration, ALSFRS-R, rate of ALSFRS-R decline, and forced vital capacity - FVC%), as well as sex and age. Disease duration, ALSFRS-R, rate of decline, and FVC% are all relevant markers of ALS severity, which were significantly correlated with each other, as expected (Fig. S2). We found no significant correlations between age and any of the disease traits or the first ten principal components,

which cumulatively explain over half of the total variance in the dataset, derived from gene expression in either vehicle or EH301 samples, suggesting that age does not significantly contribute to the variance in gene expression in our data. Sex significantly correlated with vehicle PC9 and EH301 PC1 and PC9 but did not correlate with any of the disease traits (Fig. S2). Nevertheless, we opted to include sex and age in our analysis, as they are potential biologically relevant variables in ALS. In the vehicle network, 10 modules (40%) were significantly associated with one or more traits (Fig. 3C), while in the EH301 network, 38 modules (44%) had significant associations with one or more traits (Fig. 3D).

We then performed GO pathway analyses using the GO: Biological Process (GO:BP), GO: Molecular Function (GO:MF), and KEGG databases on the significantly associated modules and found that in the vehicle network 6/10 modules (Table 2) had a significant enrichment for one or more pathways, while in the EH301 network 23/38 modules (Table 3) had a significant enrichment for one or more pathways. The module significantly associated with the largest number of traits in the vehicle network was the Greenyellow, which associated with disease status, disease duration, ALSFRS-R score, and nearly reached significant association ($p=0.06$) for FVC%. GO analysis showed that the set of genes comprising the Greenyellow module were functionally enriched for genes involved in cell cycle, chromatin modifications, and DNA damage repair (Fig. 4A). The Turquoise module significantly associated with FVC%, and neared significance for association with disease metabotype, ALSFRS-R score, and rate of decline. Genes belonging to the Turquoise module were functionally enriched for pathways related to DNA damage repair, autophagy and protein catabolism, cell cycle, innate immunity, and mitochondrial function (Fig. 4B). Interestingly, the Turquoise module was also significantly enriched for genes annotated by the KEGG database as important for ALS pathogenesis (KEGG hsa05014). Finally, the Salmon module, which significantly associated with disease metabotype, contained genes belonging to pathways related to autophagy and protein catabolism (Fig. 4C). Clustering of module eigengene expression revealed that the Salmon, Turquoise, and Greenyellow modules had highly dissimilar expression patterns from each other (Fig. S3). Furthermore, when comparing the GO terms enriched in these three modules, we found minimal overlap in terms enriched in the Greenyellow and Salmon modules (Fig. 4D), suggesting that the genes comprising them have mostly distinct functions. Although the Turquoise module has GO terms showing an over 50% overlap with those found in the Salmon and Greenyellow modules, it also has over 1000 unique GO terms, indicating that it contains genes that have functional annotations not represented in either of the other two modules. Thus, based on their eigengene expression and GO enrichment, the three modules identified in the vehicle network are non-redundant.

Table 2
Vehicle network modules with a significant trait association and GO annotation

ALS vs CTL	sALS1 vs sALS2	Disease Duration	ALSFRS-R	Rate of Decline	FVC%	Age	Sex
Greenyellow	Salmon	Greenyellow	Greenyellow		Turquoise	Brown	
						Darkred	
						Orange	

Table 3
EH301 network modules with a significant trait association and GO annotation

ALS vs CTL	sALS1 vs sALS2	Disease Duration	ALSFERS-R	Rate of Decline	FVC%	Age	Sex
Brown	White	Lightpink2	Darkviolet	Darkviolet	Red	Tan3	Salmon1
	Indianred1	Skyblue	Red	Red	Salmon1	Lavenderblush3	Navajowhite2
	Brown		Sienna4	Magenta4	Slateblue	Darkred	Slateblue
	Firebrick		Salmon1	Mediumpurple1			Magenta4
			Slateblue	Lavenderblush3			Mediumpurple1
			Brown	Lightskyblue4			Darkorange
			Magenta4				Mediumorchid4
							Magenta3

In the EH301 network, four modules were associated with at least three traits. The Brown module was associated with disease status, disease metabotype, and ALSFRS-R and included genes enriched for cell cycle, chromatin modifications, DNA damage repair, nucleic acid metabolism, and transcriptional activity GO terms (Fig. 4E). The Red module associated with ALSFRS-R, rate of decline, and FVC% and included genes enriched for chemotaxis, antigen processing, and immunity (Fig. 4F). The Slateblue module was associated with ALSFRS-R, FVC%, and sex and included genes enriched for RNA stem-loop and scaffold protein binding (not shown). Lastly, the Magenta4 module was associated with ALSFRS-R, rate of decline, and sex and included genes enriched for Fanconi anemia pathway, a pathway activated by DNA damage (not shown). Clustering of module eigengene expression revealed that Brown, Slateblue, and Magenta4 cluster together, while Red does not (Fig. S3). There was minimal overlap in GO enrichment terms among the four modules in the EH301 network that associate with traits (Fig. 4G). This indicates that, while the gene expression signatures of three of the four trait-associated modules in the EH301 network are similar, all four modules are functionally distinct.

Comparing the vehicle and EH301 networks revealed a striking difference, as the EH301 network included nearly four times the number of modules observed in vehicle treated cells. However, the intramodular connectivity was comparable between the two networks (Table 4), indicating that modules are clustered with a similar robustness in both networks. On the other hand, the total average connectivity and extramodular connectivity were significantly higher in the EH301 network (Table S4). This indicates that individual modules are more highly connected to each other in the EH301 than in the vehicle network. Modules from the two networks were compared based on their components using Fisher's exact test, and modules were paired if $p < 0.05$. This analysis revealed that 22/25 vehicle modules have a corresponding module in the EH301 network (Fig. S4). Of these 22 pairs, four significantly associated with only one common trait, while the Greenyellow and Brown pair associated with two traits common to both (disease status and ALSFRS-R). When the GO terms from all modules that significantly associated with a trait in both networks were compared, most terms associated with disease status in the Greenyellow vehicle module were also found in the Brown EH301 module (Fig. 5A), and were related to cell cycle and DNA replication (Fig. 4). This suggests that the genes that correlate with disease status likely share similar functions, regardless of treatment with EH301, but the latter introduces new associations with genes annotated with different functions, including nucleic acid metabolism and transcriptional regulation (Fig. 5A and Fig. 4E). However, for other traits including disease metabotype (Fig. 5B), disease duration (Fig. 5C), and FVC% (Fig. 5E), the overlap between GO terms from the vehicle and EH301 networks was small or absent. GO terms differing between sALS1 and sALS2 in vehicle conditions were related to autophagy and protein catabolism (Fig. 4C), while GO terms after EH301 treatment were related to glycolysis, extracellular matrix organization, cell cycle, and transcription (Supplementary Table 15). Vehicle modules associated with disease duration were enriched for terms related to cell cycle and DNA damage repair, while EH301

modules were enriched for terms related to cell adhesion, membrane polarization, and cation homeostasis (Fig. 4A and Supplementary Table 15). The Turquoise vehicle module associated with FVC% was enriched for terms related to mitochondrial function, innate immunity, cell cycle, and autophagy (Fig. 4B), while the EH301 modules associated with FVC% were enriched for terms including MHC complex assembly and antigen presentation (Fig. 4F) and cell cycle (Fig. 4E, and Supplementary Table 15). This suggests that EH301 modifies gene sets associated with disease traits, and that these genes regulate more diverse functions than those from the vehicle network. Similar to disease status (sALS or control), there was a substantial overlap in GO terms between the Greenyellow vehicle module and Brown EH301 module associated with ALSFRS-R (Fig. 5D), but there were also several modules from the EH301 network with no GO overlap with vehicle modules. These were enriched for terms including sugar alcohol metabolism, RNA and protein binding, MHC complex assembly, and cell adhesion and motility (Supplementary Table 15). Overall, comparison of the two networks revealed conservation of most of the vehicle modules after EH301 treatment, including the functionally similar Greenyellow/Brown pair that associated with disease status and ALSFRS-R. However, a large disparity was found in the functional annotation of the other modules associated with clinical traits in the two networks, indicating that EH301 modulates the expression of gene sets that are significantly correlated with various clinical traits.

Table 4
Vehicle and EH301 network structure comparison parameters

Parameter	Vehicle Network Average	EH301 Network Average	P-value
Total Connectivity	184.45	243.75	2.04 e-118
Intramodular Connectivity	140.0	144.0	0.18
Extramodular connectivity	44.0	100.0	0.00

To investigate the transcriptional regulation underlying the effects of disease and EH301 treatment on gene expression, we performed a transcription factor (TF) binding site enrichment analysis on all genes in modules significantly associated with disease traits in both networks. The TF binding sites most enriched in genes associated with disease in both vehicle and EH301 modules were those of the E2F and the Sp families of TFs (Supplementary Table 16). Importantly, enrichment of binding sites of several TFs was found only in disease-associated modules of the EH301 network, indicating that these transcriptional effects are EH301-specific (Supplementary Table 16). These TFs have also been associated with ALS in patients or model systems and include AP-2, FOXO1A [72], SREBP1 [73], MTF-1 [74, 75], and RARB [76].

Lastly, we aimed to identify “hub” genes from important modules, or genes that drive the expression profile of each module, while also correlating significantly with disease traits. Hub genes may be useful as biomarkers for potential classification of patients based on their clinical characteristics and disease severity. For each gene we calculated a significance score, denoting how strongly that gene associates with a trait, and a module membership score, denoting how closely that gene’s expression pattern matches the average module eigengene expression, or how strongly that gene “belongs” to that module [30]. We identified the most inter-connected genes in each module using connection strengths calculated from the topological overlap matrix and visualized with VisANT [77]. For each of the relevant modules in the vehicle network (Fig. S5) and the EH301 network (Fig. S6), we selected the 50 most significant genes for each associated trait, as well as the top 50 most strongly connected genes within the associated module. To illustrate the potential application of this analysis to discover disease biomarkers, we chose five genes from the Greenyellow vehicle module that were significantly associated with ALSFRS-R. To this end, genes were ranked in order of significance for correlation with ALSFRS-R and chosen if they 1) were part of the network hub identified by VisANT and 2) were identified as differentially expressed in the sALS1 vs CTL comparison in vehicle samples. Expression patterns of these five genes (*DIRAS3*, *GTSE1*, *RRM2*, *CDCA5*, and *HJURP*) showed clear differences between control and sALS1 samples, and to a lesser extent between control and sALS2 (Fig. 6A). We next examined if the expression of these genes could be used to group samples based on their ALSFRS-R score. Dendograms were constructed from hierarchical clustering first of ALSFRS-R score and then of average expression of the five genes, for vehicle-treated samples (Fig. 6B, left) and EH301-treated samples (Fig. 6B, right).

In both vehicle and EH301 treated fibroblasts, clustering based on ALSFRS-R significantly matched the clustering based on gene expression, indicating that expression of these five selected genes can be used to group samples based on their ALSFRS-R scores.

WGCNA of transcriptomic data from ALS iMNs supports and extends fibroblast results

To extend the WGCNA analysis to a cell type affected by the disease, we utilized transcriptomic and clinical data from 124 (99 ALS and 25 control) iMN lines obtained by the Answer ALS project (<https://www.answerals.org/>). We constructed a new network (Fig. 7A) using expression data from the 22,653 genes that passed quality control filters and identified 38 modules. We next calculated associations between modules and six traits (disease status, baseline ALSFRS-R, most recent ALSFRS-R, ALSFRS-R progression slope, age, and sex) and found 23 modules (60%) with a significant association with one or more traits (Fig. 7B). GO analysis revealed that 20/23 modules had significant enrichment for one or more GO:MF, GO:BP, and/or KEGG pathways (Table 5). Of these, Blue, Magenta, and Tan associated with both disease status and one or more measures of ALSFRS-R. Genes in both the Blue and Tan modules were functionally enriched for terms associated with nucleotide metabolism and several types of protein modifications, and Tan module genes were also enriched for terms involved in the mitochondrial electron transport chain (Fig. 7C, D). Genes in the Magenta module were enriched for pathways related to signal transduction through G-protein coupled receptors (Fig. S7).

Table 5
iMN network modules with a significant trait association and GO annotation

ALS	ALSFRS-R Baseline	ALSFRS-R Latest	ALSFRS-R Progression	Age	Sex
Blue	Cyan	Magenta	Blue	Honeydew1	Bisque4
Magenta		Tan	Magenta	Darkgrey	
Tan		Lightgreen	Tan	Bisque4	
		Maroon	Yellowgreen		
		Yellowgreen	Lightpink4		
		Cyan	Coral1		
		Lightpink4	Darkred		
		Darkgreen	Sienna3		
		Orangered4	Black		
		Lightyellow	Purple		
			Turquoise		

Next, we evaluated which modules in the iMN network contained genes from the Greenyellow module in the vehicle-treated fibroblast network. We found that the majority of the Greenyellow genes were found in the Purple and Turquoise modules in the iMN network, both of which significantly associate with ALSFRS-R progression slope. Accordingly, GO analysis revealed that both Purple and Turquoise genes were enriched for pathways related to cell cycle and development, similar to the pathways identified in the fibroblast Greenyellow module (Fig. 7F-G, Supplementary Table 15). We then evaluated whether the genes associated with traits in the vehicle fibroblast network had similar functional GO annotations to the genes associated with the same traits in the iMN network. There was little overlap in GO enrichment between modules associated with disease status in the fibroblast and iMN networks (Fig. 7H). However, for all traits related to ALSFRS-R, there was a large overlap in the GO terms identified in each network, with approximately 2/3 of the GO terms associated with ALSFRS-R in the fibroblast network also associating with one or more of the ALSFRS-R measures in the iMN network

(Fig. 7I). Therefore, genes related to ALSFRS-R, a measure of disease severity, share similar functions, including cell cycle and nucleic acid metabolism, in fibroblasts and iMNs.

Finally, to provide initial proof of concept of the potential predictive value of the five genes related to ALSFRS-R progression in fibroblasts (*DIRAS3*, *GTSE1*, *RRM2*, *CDCA5*, and *HJURP*), we used multinomial logistic regression to classify ALS iMN samples based on the expression of these genes. We used a ten-fold cross validation approach, in which we randomly divided the 62 ALS samples with available ALSFRS-R progression slopes into ten sets. In each iteration, one set was used as the test data and the model was trained on the remaining 9 sets. We arbitrarily defined cases as “fast progressors” if their ALSFRS-R progression slope values were one standard deviation or more below the mean. Using this method, we obtained an average accuracy of 78.9% (95% CI 66.0% - 91.6%), precision of 80.1% (95% CI 67.6% - 92.6%), false positive rate of 1.4% (95% CI -1.4% - 4.2%), and false negative rate of 17.0% (95% CI 6.3% - 28.0%), indicating good specificity and fair sensitivity. We then computed receiver operating characteristic (ROC) curves for six of the ten iterations (four iterations were unusable due to the absence of any fast progressors in the test set) and obtained an average area under the curve (AUC) value of 0.767 (95% CI 0.651 - 0.883) (Fig. 7J). This approach represents an example of how genes associated with ALSFRS-R in fibroblasts could be utilized to discriminate patients with fast disease progression relative to all other ALS cases in disease-relevant iMNs.

Discussion

Numerous lines of evidence suggest that metabolic alterations in ALS patients are related to both pathogenesis and prognosis [3–5]. Therefore, metabolism could be a viable therapeutic target for the disease. However, the biochemical and molecular underpinnings of metabolic dysregulation remain largely unknown, especially in sALS, hindering the development of effective targeted approaches to correct them. The mechanisms leading to metabolic dysregulation can differ in sALS metabolotypes, resulting in different sets of biomarkers and susceptibility to the effects of treatments. Efforts have been made to use biofluids to identify metabolic biomarkers for stratification of sALS patients [6, 7, 78], but high variability associated with environmental factors can be challenging. An alternative approach to using biofluids for unbiased -omics studies is to obtain cells from patients and investigate their metabolism under homogeneous conditions to identify specific markers indicative of sALS metabolotypes that can then be validated and in patients and used in clinical settings. Our studies have utilized primary skin fibroblasts from sALS patients to first identify a hypermetabolic phenotype relative to healthy controls [6, 7] and, more recently, to define a subtype of sALS (sALS1) metabolically characterized by enhanced transsulfuration metabolism, which was then confirmed in patient plasma [8]. This finding raised the possibility that sALS1 cases are differentially responsive to therapies that modulate cellular metabolism. The NR and PT combination drug EH301 [19] is one of these therapies, as it is designed to elevate NAD levels and activate sirtuins. EH301 was successful in a pilot ALS clinical trial [19], and is currently being investigated in a larger phase 3 clinical study. However, the metabolic effects of EH301 in different subtypes of sALS patients are still unknown.

In this study, we took advantage of a set of sALS patient-derived fibroblasts with defined metabolotype (i.e., sALS1 and sALS2) and controls, based on previous published work [8]. In these groups of cell lines, we performed in-depth transcriptomic and targeted metabolomic analyses to further elucidate profiles that characterize the subtypes and to assess how EH301 treatment affects gene expression and metabolism in the different metabolotypes. We identified striking differences in the number of DEGs, relative to control cells, between untreated sALS1 and sALS2 lines (281 vs. 1), and in the response to EH301, whereby only approximately 5% of the genes modified by the drug were in common between sALS1 and sALS2. Notably, EH301 affected a much smaller number of genes in control lines compared to both sALS groups (77 vs. >200), further indicating group-specific responses to the drug. In addition to the expected increase of NAD and its precursors, targeted metabolomics revealed that EH301 affected nucleotide metabolism in all groups, and amino acid metabolism in both sALS groups. Cysteine and methionine metabolism was only affected in sALS2, suggesting that the intrinsic transsulfuration alterations of sALS1 lines are not corrected, possibly because the effect of EH301 drives metabolism in a similar direction as the adaptive mechanisms that naturally take place in sALS1. EH301 had a potent

protective effect against metabolic stress induced by thiol group deprivation in all groups, and this effect was largely mediated by PT, because NR alone had a moderate effect only in sALS1. This result supports efficacy of EH301 in antagonizing the deleterious effects of cellular stress, but do not indicate metabotype specificity of PT. However, it needs to be noted that the stress paradigm we used is very severe and using cell death as readout may not allow for detection of more subtle phenotypic differences.

To complement and extend the differential gene expression analysis we used WGCNA, which allows association of gene expression modules with several phenotypic variables, such as clinical parameters or other indices of pathology. This approach may provide more sensitivity to detect differences in disease characteristics among patients than traditional grouped pairwise-comparison approaches [71]. WGCNA has been recently used to identify new risk genes and putative drug targets associated with major neurodegenerative diseases [79–81]. We used WGCNA to construct networks for vehicle- and EH301-treated samples and identified several modules in each network that significantly associated with disease and/or measures of disease severity. In both networks, genes belonging to modules associated with disease and ALSFRS-R were highly enriched for pathways related to the cell cycle, DNA damage repair, and nucleic acid metabolism. TF binding site analysis of disease-associated genes in both networks revealed enrichment of E2F and Sp family TF binding sites. TFs in the E2F family are essential regulators of the cell cycle and apoptosis and have been involved in the response to DNA damage in cultured neurons [82, 83]. E2F1 has been shown to have upregulated expression correlated with markers of aberrant cell cycle re-entry in postmortem spinal cord and motor cortex samples from ALS patients [84, 85]. Interestingly, in addition to its canonical role regulating the cell cycle, E2F1 is also important for controlling several aspects of global metabolic homeostasis [86]. Sp family transcription factors have been demonstrated to be key for modulating apoptosis in cultured neurons undergoing oxidative stress caused by glutathione depletion, and Sp1 knockdown is protective in G93A SOD1 mutant mice [87, 88]. Sp1 and Sp3 can also interact with E2F1 and cooperatively regulate transcription as a complex [89]. While the cell cycle has not been traditionally considered relevant in post-mitotic cells, such as motor neurons, increasing evidence suggests that cell cycle genes can be dysregulated in neurodegenerative disease. This dysregulation may promote aberrant re-entry into the cell cycle, leading to neuronal death [83, 85, 90]. Genes associated with cell cycle also have important functions in the maintenance of the cytoskeleton. Notably, mutations in kinesin genes, such as KIF5A, have been linked to ALS and other motor neuron diseases [91]. Our results confirm that cell cycle genes may be involved in ALS pathogenesis.

Despite the common effects of ALS on gene expression identified in both vehicle and EH301 networks, we also identified EH301-specific alterations. We found that EH301 markedly modifies network structure relative to vehicle, suggesting global alterations in patterns of gene expression. Evidence that EH301 produces widespread changes in transcriptional regulation is also provided by the enrichment of several TF binding sites in genes belonging to modules associated with disease, which were not found in the vehicle network. Furthermore, we identified substantial differences in the functions of genes associated with disease traits between EH301 and vehicle networks, such as cell adhesion and immune response which are unique to EH301. Therefore, EH301 may modulate genes and pathways relevant to mechanisms of ALS pathogenesis through transcriptional regulation. These findings will be valuable in interpreting the results of the ongoing clinical trial of EH301 in ALS patients.

In this study, we utilized fibroblasts as a model system, as these cells have been extensively used by us and many other groups for the investigation of molecular, biochemical, and metabolic changes in ALS. Nevertheless, we deemed it important to compare the gene expression modules obtained from ALS fibroblasts with modules from an independent, larger dataset from iMNs publicly available from the Answer ALS database. We observed a small overlap in GO terms enriched in genes belonging to modules associated with disease status in fibroblasts and iMNs, including nucleic acid metabolism terms. On the other hand, there was a substantial overlap in the functional annotations of genes associated with ALSFRS-R in both datasets, indicating disease severity-related gene expression changes common to both cell types. Previous studies have used genomic and transcriptomic data from human motor cortex to characterize sALS metabotypes, and have identified immune response, cell adhesion, cytoskeletal organization, and cell cycle among the most

dysregulated pathways in specific sALS subgroups relative to control [78, 92]. These findings coincide with our results on pathways affected in sALS fibroblasts and those identified in sALS iMNs, further supporting the notion that different subsets of sALS cases are characterized by distinct gene expression profiles that are in common among multiple cell types and tissues.

Purine and pyrimidine metabolism alterations were consistently identified in metabolomics and transcriptomics analyses in ALS fibroblasts and transcriptomics analyses in ALS iMNs. Alterations in nucleotide metabolism have also been shown in spinal cord from ALS patients [93]. In post-mitotic cells, nucleotide metabolism is required for proper DNA damage repair [94], and many of the mechanisms thought to be involved in ALS pathogenesis, such as oxidative stress and energy imbalance can lead to a loss of DNA integrity and high burden of DNA damage [95–97]. Therefore, nucleotide metabolism may be integral to the mechanisms leading to motor neuron degeneration in ALS, and nucleotide metabolism may provide new targets for therapeutic intervention.

Finally, we confirmed in iMNs the association of the five biomarker genes with ALSFRS-R that we found in fibroblasts. In agreement with the GO terms associated with disease severity identified in both fibroblasts and iMNs, four of these genes (*HJURP*, *DIRAS3*, *CDCA5*, and *GTSE1*) are associated with cell cycle and DNA damage repair, while *RRM2* is involved in purine and pyrimidine metabolism. To establish an initial proof of concept that the expression of a small subset of genes in patient-derived cells could be used to predict disease progression, we employed a logistic regression model based on the expression of the five biomarker genes in iMNs to identify fast progressing sALS cases. Cross-validation of ROC results showed an average accuracy of 78.9% and precision of 80.1%, which indicates that the combination of the expression of these genes could potentially identify fast progressing sALS cases. This proof-of-concept assessment was performed on a relatively small dataset available from Answer ALS, but we think that the encouraging results warrant the extension of these studies to larger datasets from patient-derived cells, when they become available.

Conclusions

Multi-omics analyses of patient-derived fibroblasts highlight differential metabolic and transcriptomic profiles in sALS metabolotypes, which translate into differential responses to the investigational drug EH301. In the future, it will be important to apply similar multi-omics and machine learning strategies to readily available patient materials, such as white blood cells, to develop viable predictive biomarkers of disease progression and response to EH301 and other therapeutics.

Abbreviations

sALS, sporadic amyotrophic lateral sclerosis; ALSFRS-R, revised ALS functional rating scale; FVC, forced vital capacity; GO, gene ontology; GSH, glutathione; iMN, induced motor neurons; NAD, nicotinamide adenine dinucleotide; NR, nicotinamide riboside; PT, pterostilbene; TF, transcription factor; WGCNA, weighted gene co-expression network analysis.

Declarations

Ethics approval and consent to participate

All human samples were obtained from repositories of de-identified cell lines. The academic centers that provided cell lines had approved IRB protocols in place.

Consent for publication

Not applicable

Availability of data and materials

All R packages are available from CRAN. Full differential expression and gene ontology analysis results are included in Supplementary Tables. Raw transcriptomic data will be made available from Gene Expression Omnibus. Transcriptomic data from induced motor neurons are available from Answer ALS.

Competing interests

The authors declare that they have no financial or non-financial competing interests.

Funding

This work was supported by funds from ALS Association grant ALS-90452 (to GM), NIH/NINDS grant R35 NS122209-01 (to GM), and Muscular Dystrophy Association grant 602762 (to CK).

Authors' contributions

JAF and GC designed and performed experiments, analyzed data, and wrote the manuscript. CK performed experiments and analyzed data. HH and RWD conceived experiments and wrote the manuscript. GM conceived and designed experiments and wrote the manuscript. All authors read and approved the manuscript prior to submission.

Acknowledgements

We thank Elysium Health for providing NR and PT for this study. We thank Dr. Hiroshi Mitsumoto (Columbia University) and the COSMOS initiative for providing the fibroblast and plasma samples utilized in this work. Fibroblast samples and clinical data were obtained from the NINDS Repository.

References

1. Turner MR, et al. Genetic screening in sporadic ALS and FTD. *J Neurol Neurosurg Psychiatry*. 2017;88(12):1042–4.
2. Kiernan MC, et al. Improving clinical trial outcomes in amyotrophic lateral sclerosis. *Nat Rev Neurol*. 2021;17(2):104–18.
3. Huisman MH, et al. Effect of Presymptomatic Body Mass Index and Consumption of Fat and Alcohol on Amyotrophic Lateral Sclerosis. *JAMA Neurol*. 2015;72(10):1155–62.
4. Cistaro A, et al. Brain hypermetabolism in amyotrophic lateral sclerosis: a FDG PET study in ALS of spinal and bulbar onset. *Eur J Nucl Med Mol Imaging*. 2012;39(2):251–9.
5. Steyn FJ, et al. Hypermetabolism in ALS is associated with greater functional decline and shorter survival. *J Neurol Neurosurg Psychiatry*. 2018;89(10):1016–23.
6. Kirk K, et al. Bioenergetic markers in skin fibroblasts of sporadic amyotrophic lateral sclerosis and progressive lateral sclerosis patients. *Ann Neurol*. 2014;76(4):620–4.
7. Konrad C, et al. Fibroblast bioenergetics to classify amyotrophic lateral sclerosis patients. *Mol Neurodegener*. 2017;12(1):76.
8. Chen Q, et al. Accelerated transsulfuration metabolically defines a discrete subclass of amyotrophic lateral sclerosis patients. *Neurobiol Dis*. 2020;144:105025.
9. Belenky P, Bogan KL, Brenner C. NAD⁺ metabolism in health and disease. *Trends Biochem Sci*. 2007;32(1):12–9.
10. Massudi H, et al. Age-associated changes in oxidative stress and NAD⁺ metabolism in human tissue. *PLoS One*. 2012;7(7):e42357.
11. Harlan BA, et al. Evaluation of the NAD(+) biosynthetic pathway in ALS patients and effect of modulating NAD(+) levels in hSOD1-linked ALS mouse models. *Exp Neurol*. 2020;327:113219.

12. Harlan BA, et al. Enhancing NAD⁺ Salvage Pathway Reverts the Toxicity of Primary Astrocytes Expressing Amyotrophic Lateral Sclerosis-linked Mutant Superoxide Dismutase 1 (SOD1). *J Biol Chem*. 2016;291(20):10836–46.
13. Dellinger RW, et al. Repeat dose NRPT (nicotinamide riboside and pterostilbene) increases NAD(+) levels in humans safely and sustainably: a randomized, double-blind, placebo-controlled study. *NPJ Aging Mech Dis*. 2017;3:17.
14. Li YR, Li S, Lin CC. Effect of resveratrol and pterostilbene on aging and longevity. *Biofactors*. 2018;44(1):69–82.
15. Wang P, Sang S. Metabolism and pharmacokinetics of resveratrol and pterostilbene. *Biofactors*. 2018;44(1):16–25.
16. Gomez-Zorita S, et al. Effects of Pterostilbene on Diabetes, Liver Steatosis and Serum Lipids. *Curr Med Chem*. 2021;28(2):238–52.
17. Freysson A, et al. Natural stilbenes effects in animal models of Alzheimer's disease. *Neural Regen Res*. 2020;15(5):843–9.
18. Obrador E, et al. Nicotinamide Riboside and Pterostilbene Cooperatively Delay Motor Neuron Failure in ALS SOD1(G93A) Mice. *Mol Neurobiol*; 2020.
19. de la Rubia JE, et al. Efficacy and tolerability of EH301 for amyotrophic lateral sclerosis: a randomized, double-blind, placebo-controlled human pilot study. *Amyotroph Lateral Scler Frontotemporal Degener*. 2019;20(1-2):115–22.
20. NO-ALS. *The NO-ALS Study: A Trial of Nicotinamide/Pterostilbene Supplement in ALS*. Available from: <https://clinicaltrials.gov/ct2/show/NCT04562831?term=eh301&cond=ALS&draw=2&rank=2>.
21. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
22. Dobin A, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics*. 2013;29(1):15–21.
23. Li H, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics*. 2009;25(16):2078–9.
24. Anders S, Pyl PT, Huber W. HTSeq—a Python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31(2):166–9.
25. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
26. Liao Y, et al. WebGestalt 2019: gene set analysis toolkit with revamped UIs and APIs. *Nucleic Acids Res*. 2019;47(W1):W199–205.
27. Xia J, et al., *MetaboAnalyst: a web server for metabolomic data analysis and interpretation*. *Nucleic Acids Res*, 2009. **37**(Web Server issue): p. W652-60.
28. Raudvere U, et al. g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res*. 2019;47(W1):W191–8.
29. Granatiero V, et al., *Modulation of the IGF1R-MTOR pathway attenuates motor neuron toxicity of human ALS SOD1(G93A) astrocytes*. *Autophagy*, 2021: p. 1–14.
30. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
31. Langfelder P, et al. Is my network module preserved and reproducible? *PLoS Comput Biol*. 2011;7(1):e1001057.
32. Ritchie ME, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
33. Galili T. dendextend: an R package for visualizing, adjusting and comparing trees of hierarchical clustering. *Bioinformatics*. 2015;31(22):3718–20.
34. Baker FB, Hubert LJ. Measuring the Power of Hierarchical Cluster Analysis. *J Am Stat Assoc*. 1975;70(349):31–8.
35. Kuhn M. *The caret Package*. *Journal of Statistical Software*, 2012. 28.
36. Sing T, et al. ROCR: visualizing classifier performance in R. *Bioinformatics*. 2005;21(20):3940–1.

37. Klim JR, et al. ALS-implicated protein TDP-43 sustains levels of STMN2, a mediator of motor neuron growth and repair. *Nat Neurosci.* 2019;22(2):167–79.
38. Theunissen F, et al. Novel STMN2 Variant Linked to Amyotrophic Lateral Sclerosis Risk and Clinical Phenotype. *Front Aging Neurosci.* 2021;13:658226.
39. Kanai Y, et al. KIF5C, a novel neuronal kinesin enriched in motor neurons. *J Neurosci.* 2000;20(17):6374–84.
40. Padzik A, et al. KIF5C S176 Phosphorylation Regulates Microtubule Binding and Transport Efficiency in Mammalian Neurons. *Front Cell Neurosci.* 2016;10:57.
41. Poirier K, et al. Mutations in TUBG1, DYNC1H1, KIF5C and KIF2A cause malformations of cortical development and microcephaly. *Nat Genet.* 2013;45(6):639–47.
42. Michels S, et al. Mutations of KIF5C cause a neurodevelopmental disorder of infantile-onset epilepsy, absent language, and distinctive malformations of cortical development. *Am J Med Genet A.* 2017;173(12):3127–31.
43. Nishimura AL, et al. A mutation in the vesicle-trafficking protein VAPB causes late-onset spinal muscular atrophy and amyotrophic lateral sclerosis. *Am J Hum Genet.* 2004;75(5):822–31.
44. Kuijpers M, et al. The ALS8 protein VAPB interacts with the ER-Golgi recycling protein YIF1A and regulates membrane delivery into dendrites. *EMBO J.* 2013;32(14):2056–72.
45. Tessadori F, et al. Germline mutations affecting the histone H4 core cause a developmental syndrome by altering DNA damage response and cell cycle control. *Nat Genet.* 2017;49(11):1642–6.
46. Byun JS, et al. The transcription factor PITX1 drives astrocyte differentiation by regulating the SOX9 gene. *J Biol Chem.* 2020;295(39):13677–90.
47. Sun W, et al. SOX9 Is an Astrocyte-Specific Nuclear Marker in the Adult Brain Outside the Neurogenic Regions. *J Neurosci.* 2017;37(17):4493–507.
48. Zhang Q, et al. Nesprin-2 is a multi-isomeric protein that binds lamin and emerin at the nuclear envelope and forms a subcellular network in skeletal muscle. *J Cell Sci.* 2005;118(Pt 4):673–87.
49. Zhang X, et al. Syne-1 and Syne-2 play crucial roles in myonuclear anchorage and motor neuron innervation. *Development.* 2007;134(5):901–8.
50. Rahman S, et al., *A Survey of Reported Disease-Related Mutations in the MRE11-RAD50-NBS1 Complex.* *Cells*, 2020. 9(7).
51. Rizzo F, et al. Key role of SMN/SYNCRIP and RNA-Motif 7 in spinal muscular atrophy: RNA-Seq and motif analysis of human motor neurons. *Brain.* 2019;142(2):276–94.
52. Lelieveld SH, et al. Meta-analysis of 2,104 trios provides support for 10 new genes for intellectual disability. *Nat Neurosci.* 2016;19(9):1194–6.
53. Gillentine MA, et al. Rare deleterious mutations of HNRNP genes result in shared neurodevelopmental disorders. *Genome Med.* 2021;13(1):63.
54. Busetto V, et al. Structural and functional insights into CWC27/CWC22 heterodimer linking the exon junction complex to spliceosomes. *Nucleic Acids Res.* 2020;48(10):5670–83.
55. Moretti J, et al. The translation initiation factor 3f (eIF3f) exhibits a deubiquitinase activity regulating Notch activation. *PLoS Biol.* 2010;8(11):e1000545.
56. Elantak L, et al. The indispensable N-terminal half of eIF3j/HCR1 cooperates with its structurally conserved binding partner eIF3b/PRT1-RRM and with eIF1A in stringent AUG selection. *J Mol Biol.* 2010;396(4):1097–116.
57. Young DJ, Guydosh NR. Hcr1/eIF3j Is a 60S Ribosomal Subunit Recycling Accessory Factor In Vivo. *Cell Rep.* 2019;28(1):39–50 e4.
58. Nickel W, Rabouille C. Mechanisms of regulated unconventional protein secretion. *Nat Rev Mol Cell Biol.* 2009;10(2):148–55.

59. Trychta KA, et al. KDEL Receptors Are Differentially Regulated to Maintain the ER Proteome under Calcium Deficiency. *Cell Rep.* 2018;25(7):1829–40 e6.
60. Annaert WG, et al. Export of cellubrevin from the endoplasmic reticulum is controlled by BAP31. *J Cell Biol.* 1997;139(6):1397–410.
61. Wakana Y, et al. Bap31 is an itinerant protein that moves between the peripheral endoplasmic reticulum (ER) and a juxtannuclear compartment related to ER-associated Degradation. *Mol Biol Cell.* 2008;19(5):1825–36.
62. Namba T. BAP31 regulates mitochondrial function via interaction with Tom40 within ER-mitochondria contact sites. *Sci Adv.* 2019;5(6):eaaw1386.
63. Jin H, Komita M, Aoe T. The Role of BiP Retrieval by the KDEL Receptor in the Early Secretory Pathway and its Effect on Protein Quality Control and Neurodegeneration. *Front Mol Neurosci.* 2017;10:222.
64. Whalen S, et al., *Further delineation of BCAP31-linked intellectual disability: description of 17 new families with LoF and missense variants.* *Eur J Hum Genet*, 2021.
65. Dominissini D, et al. The dynamic N(1)-methyladenosine methylome in eukaryotic messenger RNA. *Nature.* 2016;530(7591):441–6.
66. Zhang C, Jia G. Reversible RNA Modification N(1)-methyladenosine (m(1)A) in mRNA and tRNA. *Genomics Proteomics Bioinformatics.* 2018;16(3):155–61.
67. Li X, et al. Base-Resolution Mapping Reveals Distinct m(1)A Methylome in Nuclear- and Mitochondrial-Encoded Transcripts. *Mol Cell.* 2017;68(5):993–1005 e9.
68. Sattler R, Rothstein JD. *Regulation and dysregulation of glutamate transporters.* *Handb Exp Pharmacol*, 2006(175): p. 277–303.
69. Sears TK, Angelastro JM. The transcription factor ATF5: role in cellular differentiation, stress responses, and cancer. *Oncotarget.* 2017;8(48):84595–609.
70. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol.* 2005;4:Article17.
71. Gaiteri C, et al. Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders. *Genes Brain Behav.* 2014;13(1):13–24.
72. Bernardini C, et al. Mitochondrial network genes in the skeletal muscle of amyotrophic lateral sclerosis patients. *PLoS One.* 2013;8(2):e57739.
73. Taghibiglou C, et al. Sterol regulatory element binding protein-1 (SREBP1) activation in motor neurons in excitotoxicity and amyotrophic lateral sclerosis (ALS): Indip, a potential therapeutic peptide. *Biochem Biophys Res Commun.* 2011;413(2):159–63.
74. Dangoumau A, et al., *Dysregulations of Expression of Genes of the Ubiquitin/SUMO Pathways in an In Vitro Model of Amyotrophic Lateral Sclerosis Combining Oxidative Stress and SOD1 Gene Mutation.* *Int J Mol Sci*, 2021. 22(4).
75. Morahan JM, et al. Genetic susceptibility to environmental toxicants in ALS. *Am J Med Genet B Neuropsychiatr Genet.* 2007;144B(7):885–90.
76. Kim KY, et al. Downregulated miR-18b-5p triggers apoptosis by inhibition of calcium signaling and neuronal cell differentiation in transgenic SOD1 (G93A) mice and SOD1 (G17S and G86S) ALS patients. *Transl Neurodegener.* 2020;9(1):23.
77. Hu Z, Snitkin ES, DeLisi C. VisANT: an integrative framework for networks in systems biology. *Brief Bioinform.* 2008;9(4):317–25.
78. Aronica E, et al. Molecular classification of amyotrophic lateral sclerosis by unsupervised clustering of gene expression in motor cortex. *Neurobiol Dis.* 2015;74:359–76.

79. Wang JC, Ramaswami G, Geschwind DH. Gene co-expression network analysis in human spinal cord highlights mechanisms underlying amyotrophic lateral sclerosis susceptibility. *Sci Rep.* 2021;11(1):5748.
80. Mostafavi S, et al. A molecular network of the aging human brain provides insights into the pathology and cognitive decline of Alzheimer's disease. *Nat Neurosci.* 2018;21(6):811–9.
81. Sun Y, Lin J, Zhang L. The application of weighted gene co-expression network analysis in identifying key modules and hub genes associated with disease status in Alzheimer's disease. *Ann Transl Med.* 2019;7(24):800.
82. Park DS, et al. Cyclin-dependent kinases participate in death of neurons evoked by DNA-damaging agents. *J Cell Biol.* 1998;143(2):457–67.
83. Zhang Y, Song X, Herrup K. Context-Dependent Functions of E2F1: Cell Cycle, Cell Death, and DNA Damage Repair in Cortical Neurons. *Mol Neurobiol.* 2020;57(5):2377–90.
84. Ranganathan S, Bowser R. Alterations in G(1) to S phase cell-cycle regulators during amyotrophic lateral sclerosis. *Am J Pathol.* 2003;162(3):823–35.
85. Joseph C, et al. Cell Cycle Deficits in Neurodegenerative Disorders: Uncovering Molecular Mechanisms to Drive Innovative Therapeutic Development. *Aging Dis.* 2020;11(4):946–66.
86. Denechaud PD, Fajas L, Giralt A. E2F1, a Novel Regulator of Metabolism. *Front Endocrinol (Lausanne).* 2017;8:311.
87. Ryu H, et al. Sp1 and Sp3 are oxidative stress-inducible, antideath transcription factors in cortical neurons. *J Neurosci.* 2003;23(9):3597–606.
88. Garcia-Morales V, et al. Sp1-regulated expression of p11 contributes to motor neuron degeneration by membrane insertion of TASK1. *Nat Commun.* 2019;10(1):3784.
89. Rotheneder H, Geymayer S, Haidweger E. Transcription factors of the Sp1 family: interaction with E2F and regulation of the murine thymidine kinase promoter. *J Mol Biol.* 1999;293(5):1005–15.
90. Zhang X, Song S, Peng W. *Cell cycle deregulation in neurodegenerative diseases.* *Int J Neurosci,* 2021: p. 1–15.
91. Nicolas A, et al. Genome-wide Analyses Identify KIF5A as a Novel ALS Gene. *Neuron.* 2018;97(6):1268–83 e6.
92. Morello G, et al. Integrative multi-omic analysis identifies new drivers and pathways in molecularly distinct subtypes of ALS. *Sci Rep.* 2019;9(1):9968.
93. Veyrat-Durebex C, et al. Metabo-lipidomics of Fibroblasts and Mitochondrial-Endoplasmic Reticulum Extracts from ALS Patients Shows Alterations in Purine, Pyrimidine, Energetic, and Phospholipid Metabolisms. *Mol Neurobiol.* 2019;56(8):5780–91.
94. Fishel ML, Vasko MR, Kelley MR. DNA repair in neurons: so if they don't divide what's to repair? *Mutat Res.* 2007;614(1-2):24–36.
95. Penndorf D, Witte OW, Kretz A. DNA plasticity and damage in amyotrophic lateral sclerosis. *Neural Regen Res.* 2018;13(2):173–80.
96. Schwartz EI, et al. Cell cycle activation in postmitotic neurons is essential for DNA repair. *Cell Cycle.* 2007;6(3):318–29.
97. Coppede F. An overview of DNA repair in amyotrophic lateral sclerosis. *ScientificWorldJournal.* 2011;11:1679–91.

Figures

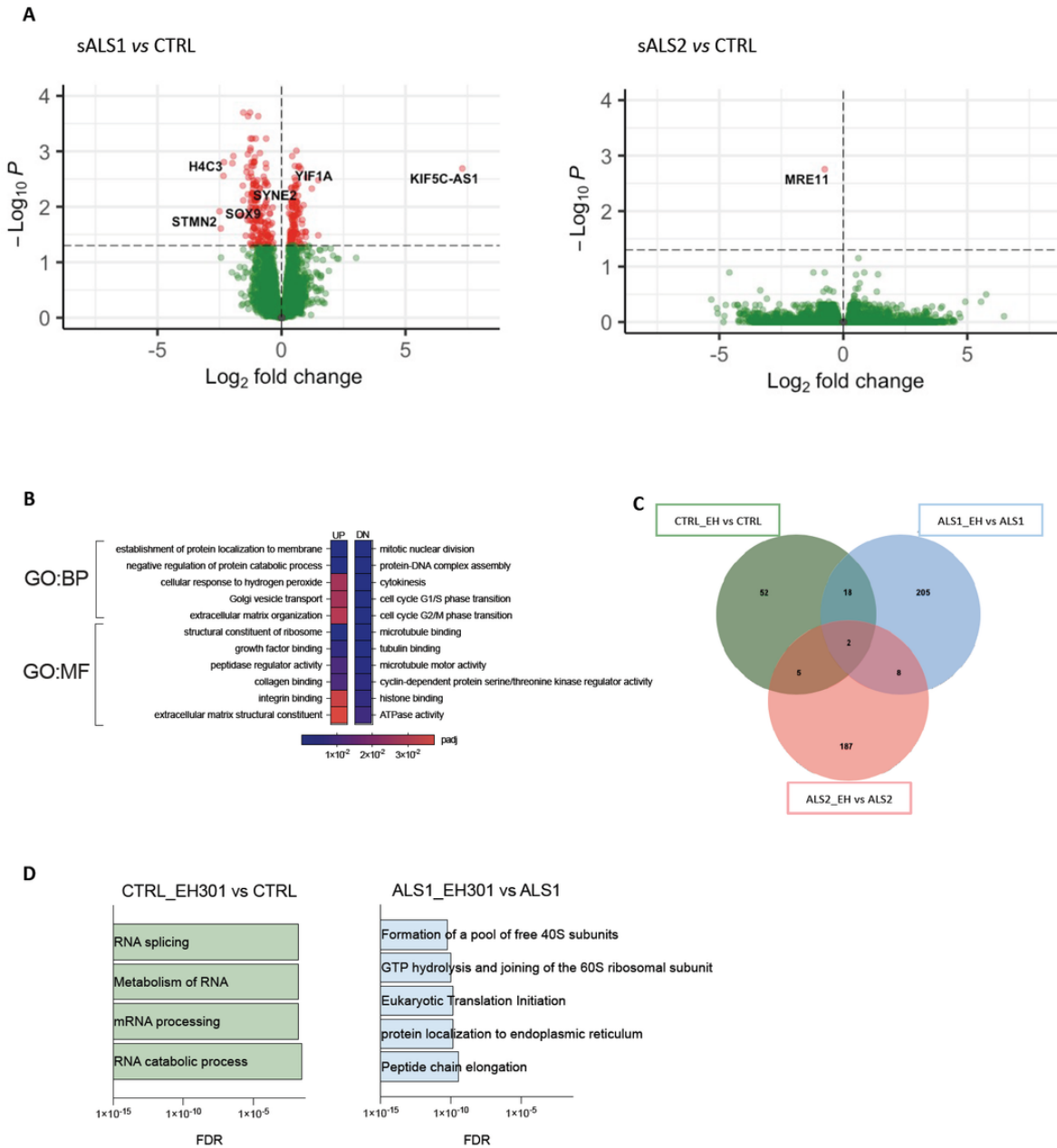


Figure 1

Different sALS metabolotypes show distinct gene expression profiles and transcriptional responses to EH301. a) Volcano plot of DEGs in sALS1 (left) versus control and sALS2 (right) versus control fibroblasts (n=6 per group). Red dots represent DEGs with p value < 0.05 after FDR adjustment. Known ALS-related disease genes are labeled. b) Gene ontology pathways with significantly enriched biological processes (BP) and molecular functions (MF) of upregulated (UP) and downregulated (DN) DEGs in sALS1 compared to control. c) Venn diagram of DEGs with p value < 0.05 after FDR adjustment modified by EH301 in control (green), sALS1 (blue), and sALS2 (red) fibroblasts (n=6 per group). d) Gene ontology pathways significantly enriched in EH301 treated controls (green) and sALS1 (blue) fibroblasts, (n=6 per group), all DEGs with p value < 0.05 after FDR adjustment were included in the analysis.

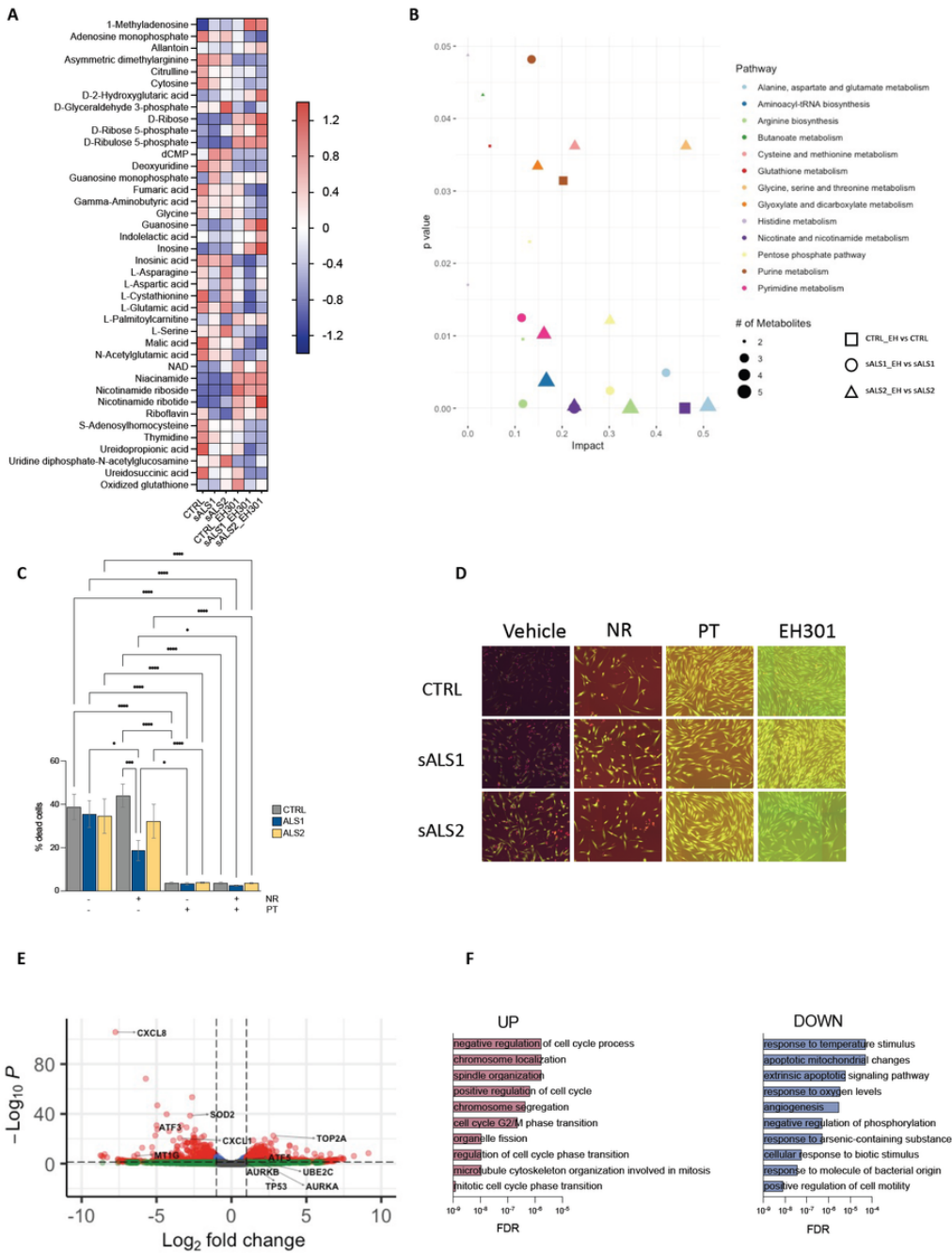


Figure 2

EH301 affects nucleotide and amino acid metabolism and protects fibroblasts from thiol depletion stress. a) Z score heatmap of metabolites significantly modified by EH301 in control, sALS1, and sALS2 fibroblasts (n=6 per group), p value < 0.05 after FDR adjustment. b) Scatter plot of metabolic pathways enriched in control (square), sALS1 (circle), and sALS2 (triangle) fibroblasts after treatment with EH301, shape size reflects number of metabolites. c) Percentage cell death in control, sALS1, and sALS2 fibroblasts grown in methionine/cystine depleted medium for 72h and treated with EH301 or its single components (NR, PT). Bars represent average +/- standard error of mean of at least three independent experiments. One-way ANOVA followed by Fisher LSD. *p<0.05, **p<0.01, ***p<0.0001. d) Representative images of control, sALS1, and sALS2 fibroblasts stained with DAPI, calcein and propidium iodide to assess cell viability. e) Volcano plot of transcripts modified by EH301 in control fibroblasts (n=3 per condition) grown for 48h in methionine/cystine depleted medium compared to vehicle treatment. Red dots represent DEGs with log2 fold change greater than 1 with p value < 0.05 after FDR adjustment. Representative genes belonging to enriched GO pathways are labeled. f) Significantly enriched pathways of

upregulated (UP) and downregulated (DOWN) genes in control fibroblasts grown in methionine/cystine depleted medium for 48h and treated with EH301 (n=3 per condition).

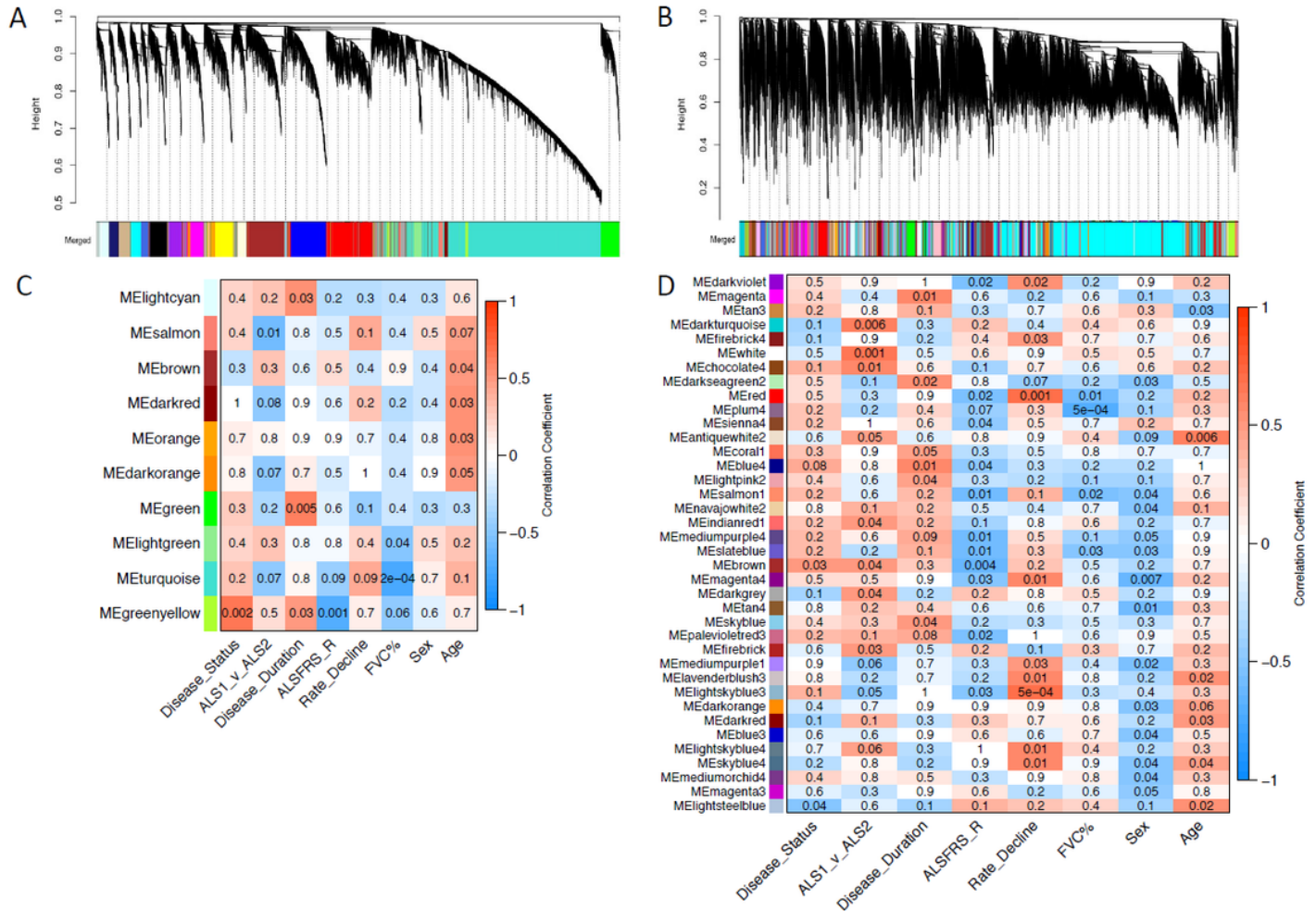


Figure 3

EH301 modifies network structure and associations of modules with disease traits. a-b) Dendrogram of hierarchical clustering of gene co-expression dissimilarity values, constructed from the topological overlap matrix, for vehicle-treated samples (a) and EH301-treated samples (b). Colors correspond to module assignments (see Fig S1 for network construction parameters). c-d) Heatmaps showing correlations between module eigengene expression values and clinical traits for vehicle-treated samples (c) and EH301-treated samples (d). The numbers in each box are p-values, while box colors correspond to the correlation coefficient. For clarity, only modules with at least 1 significant trait association are shown.

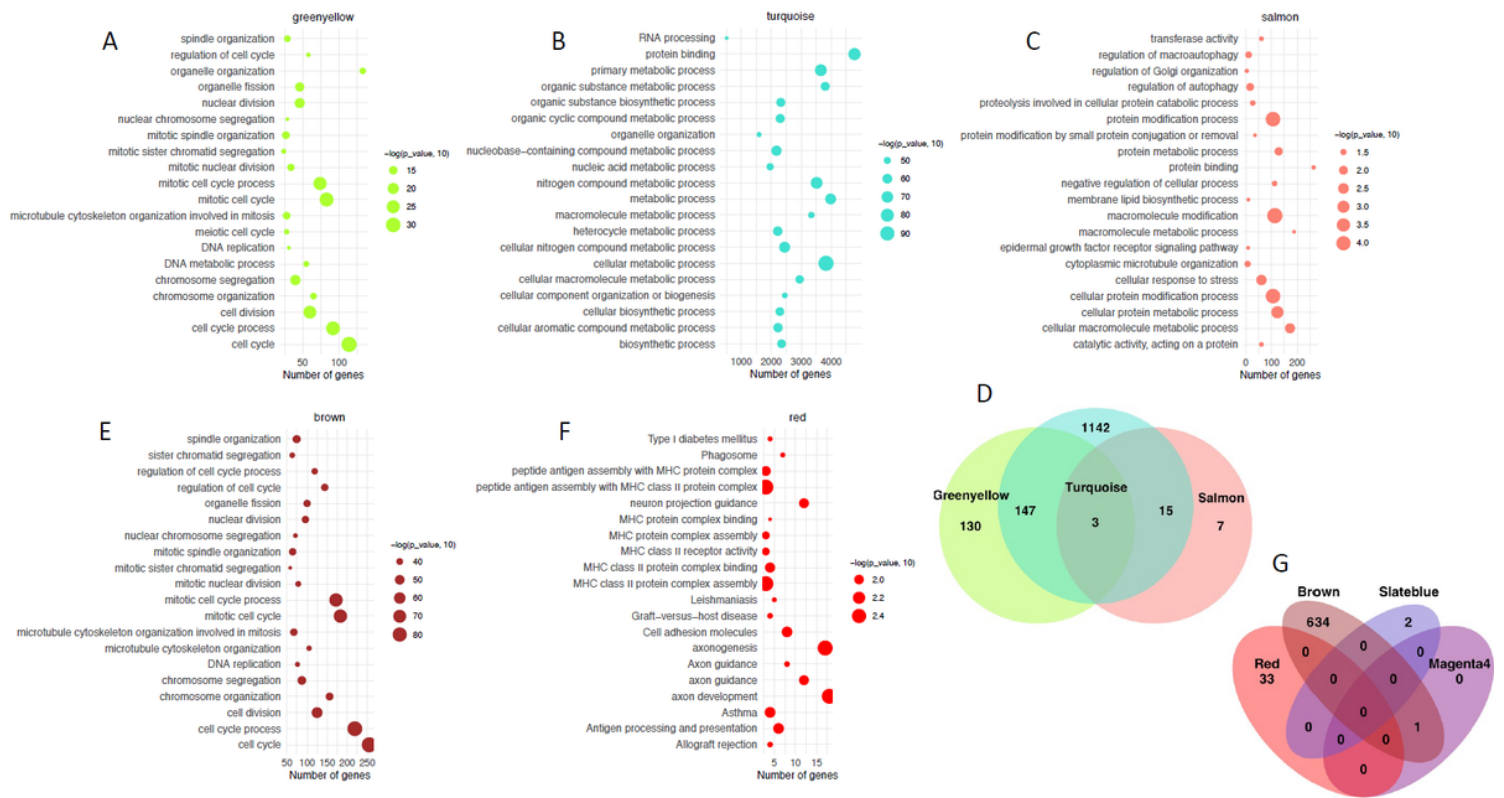


Figure 4

Both vehicle and EH301 disease-associated modules have overlapping and unique GO functional annotations. a-c) Top 20 most significant GO terms enriched in the Greenyellow (a), Turquoise (b), and Salmon (c) modules from the vehicle network. d) Venn Diagram showing overlap of all significantly enriched GO terms in the Greenyellow, Turquoise, and Salmon vehicle modules. e-f) Top 20 most significant GO terms enriched in the Brown (e) and Red (f) modules from the EH301 network. g. Venn Diagram showing overlap of all significantly enriched GO terms in the Brown, Red, Slateblue, and Magenta4 EH301 modules.

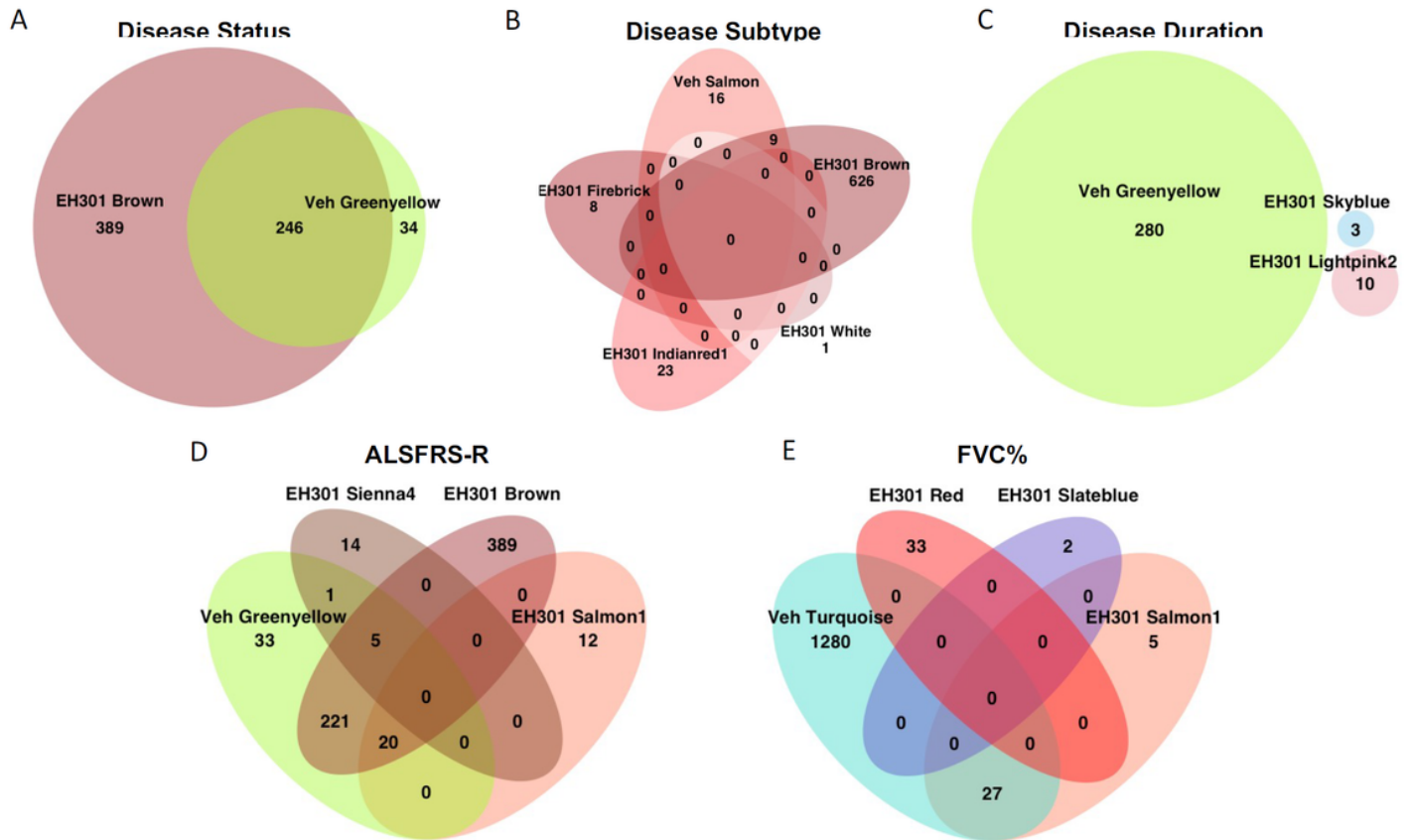


Figure 5

Minimal overlap exists between GO terms from vehicle and EH301 modules associated with specific disease traits. a-e) Venn Diagrams comparing overlap between all significantly enriched GO terms from all modules significantly associated with the listed clinical traits, from vehicle and EH301 networks. Due to the number of EH301 modules associated with ALSFRS-R score, only the 4 modules containing common terms are shown.

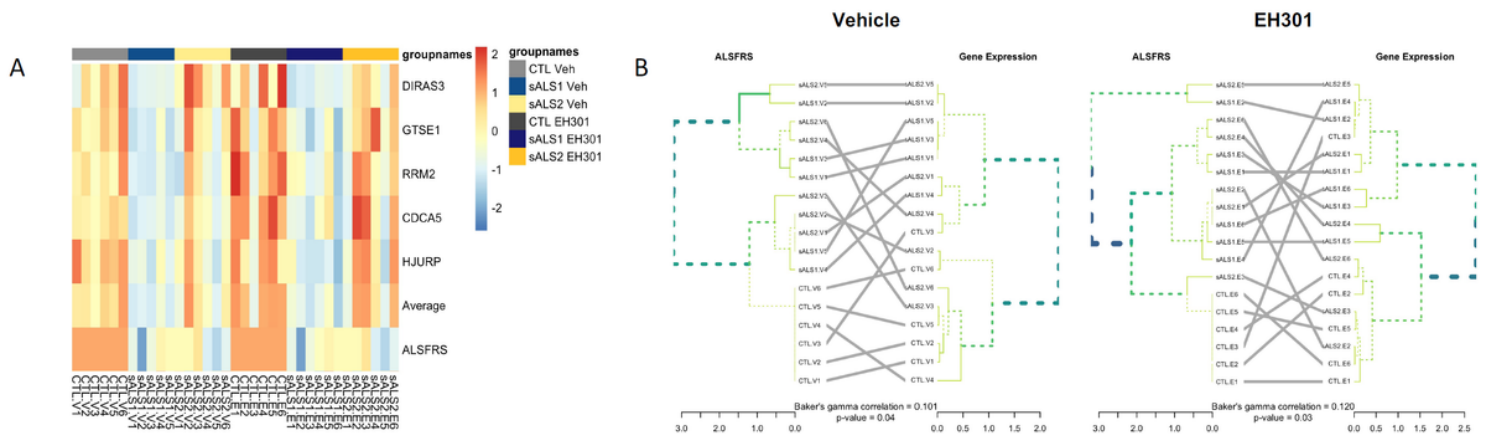


Figure 6

A set of five hub genes that correlate strongly with ALSFRS-R can be used to cluster samples into groups that correspond to their ALSFRS-R scores. a) Heatmap showing Z-scores of normalized expression values of the five genes chosen as “markers” of ALSFRS-R score, along with aggregated average expression values of all five and ALSFRS-R score for each

sample. b) Dendrograms created by hierarchical clustering of ALSFRS-R score on the left, and average expression of the five marker genes on the right, for vehicle samples (left) and EH301 samples (right). Central gray lines show matching of each sample over the two trees, and colored lines in the tree denote matching clusters. P-value for Baker's gamma index calculated using 100-fold permutation.

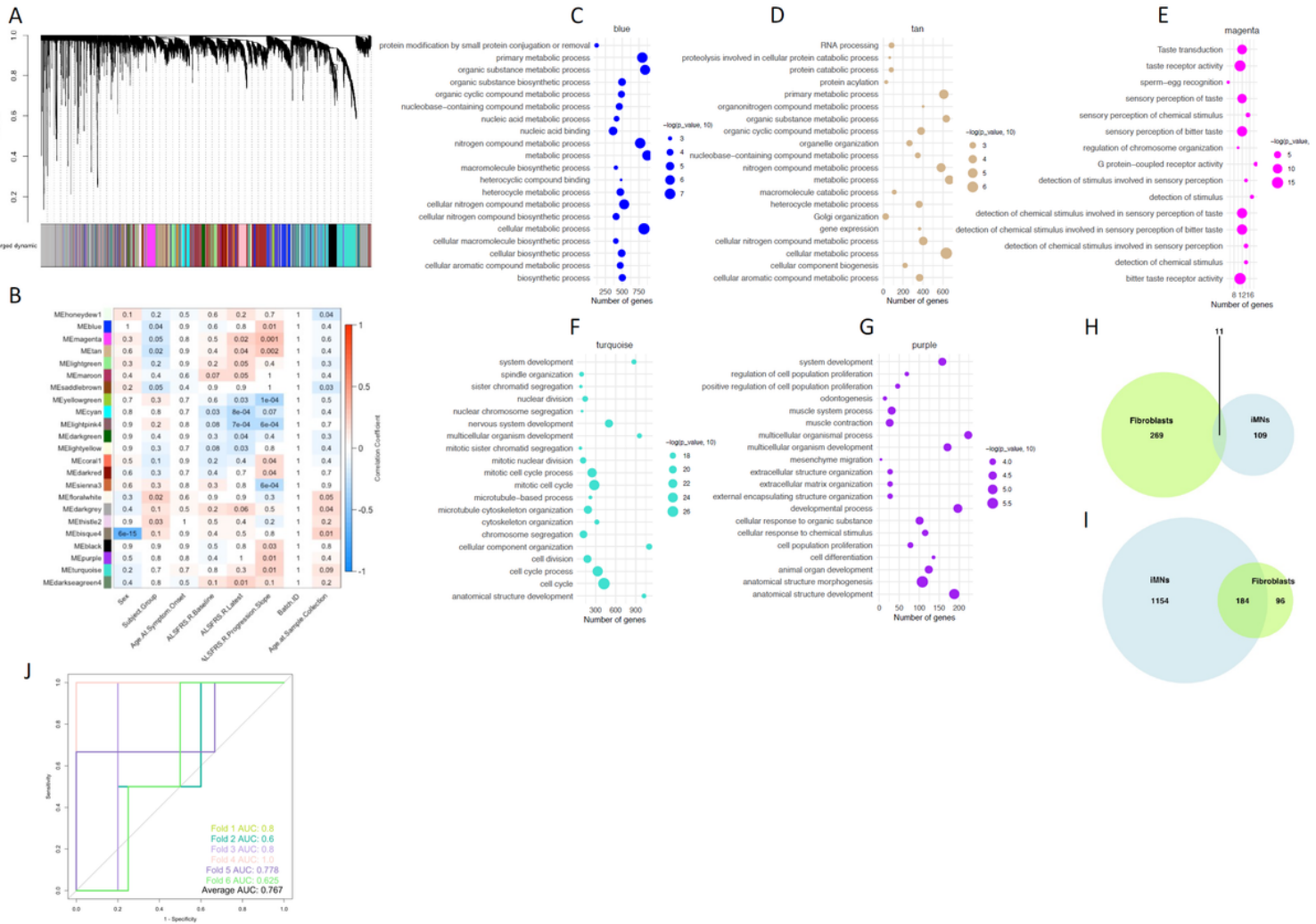


Figure 7

WGCNA of transcriptomic data from iMNs confirms the association of several GO pathways with disease traits found in fibroblasts. a) Dendrogram of hierarchical clustering of gene co-expression dissimilarity values, constructed from the topological overlap matrix, for all iMN samples (n= 124). Colors correspond to module assignments. b) Heatmap showing correlations between module eigengene expression values and clinical traits. The numbers in each box are p-values, while box colors correspond to the correlation coefficient. For clarity, only modules with at least 1 significant trait association are shown. c-g. Top 20 most significant GO terms enriched in the Blue (c), Tan (d), Magenta (e), Turquoise (f), and Purple (g) modules. h-i. Venn Diagram showing overlap of all significantly enriched GO terms from all modules associated with disease status (h) or ALSFRS-R (i) in the iMN and vehicle fibroblast networks. j) ROC curves for each of the six cross-validation runs with computable ROC. AUC values for each curve indicated on bottom right.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [FigureS1.pdf](#)
- [FigureS2.pdf](#)
- [FigureS3.pdf](#)
- [FigureS4.pdf](#)
- [FigureS5.pdf](#)
- [FigureS6.pdf](#)
- [SupplementaryTables.xlsx](#)