Transcriptional modulation unique to vulnerable motor neurons predicts ALS across species and SOD1 mutations

Irene Mei[1], Susanne Nichterwitz[1,3], Melanie Leboeuf[1,2], Jik Nijssen[2,3], Isadora Lenoel[4], Dirk Repsilber[5], Christian S. Lobsiger[4], Eva Hedlund[1,2,3]

[1] Department of Biochemistry and Biophysics, Stockholm University, Stockholm, Sweden

[2] Department of Cellular and Molecular Biology, Karolinska Institutet, Stockholm, Sweden

[3] Department of Neuroscience, Karolinska Institutet, Stockholm, Sweden

[4] Sorbonne Université, Institut du Cerveau-Paris Brain Institute - ICM, Inserm, CNRS, Paris, France

[5] School of Medical Sciences, Örebro University, Örebro, Sweden

Corresponding author: eva.hedlund@dbb.su.se

Abstract

Amyotrophic lateral sclerosis (ALS) is characterized by the progressive loss of somatic motor neurons (MNs), which innervate skeletal muscles. However, certain MN groups including ocular MNs that regulate eye movement are relatively resilient to ALS. To reveal mechanisms of differential susceptibility, we investigate the transcriptional dynamics of two vulnerable and two resilient MN populations in SOD1G93A ALS mice. Analysis of differentially expressed genes (DEGs) shows that each neuron type displays a largely unique spatial and temporal response to disease. Baseline gene expression of resilient MNs is clearly divergent from vulnerable MNs and their response to mutant SOD1 is minor, with regulation of few genes, some with known neuroprotective properties, including e.g. *Ucn*, *Cck* and *Postn*. EASE, fGSEA and ANUBIX enrichment analysis demonstrate that vulnerable MN groups share pathway activation, including regulation of neuronal death, inflammatory response, ERK and MAPK cascades, cell adhesion and synaptic signaling. These pathways are largely driven by 11 upregulated genes, including *Atf3, Nupr1, Cd44, Gadd45a, Ngfr, Ccl2, Ccl7, Gal, Timp1, Serpinb1a* and *Chl1*, and indicate that cell death occurs through similar mechanisms across vulnerable MNs albeit with distinct timing. Machine learning using DEGs upregulated in our SOD1G93A spinal MNs predict disease in human stem cell-derived MNs harboring the *SOD1E100G* mutation, and show that dysregulation of *VGF*, *INA, PENK* and *NTS* are strong disease-predictors across SOD1 mutations and species. Meta-analysis across SOD1 transcriptome datasets identifies a shared vulnerability code of 32 genes including e.g *Atf3*, *Nupr1*, *Vgf*, *Ina*, *Sprr1a, Fgf21, C1qb, Gap43, Adcyap1,* and *Mt1*. In conclusion our study reveals cell-type-specific gene expression and dynamic disease-induced regulation that may act to preserve neurons and can be used to predict disease.

Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease that is characterized by the loss of upper motor neurons in the cortex, leading to spasticity, and lower somatic motor neurons in the brainstem and spinal cord that control skeletal muscles, leading to muscle atrophy and paralysis. About 10% of cases are familial and of these, around 20% are caused by dominant mutations in the superoxide dismutase 1 (SOD1) gene, which induce ALS through acquired gain-of-function mechanisms (Rosen et al. 1993). SOD1 is ubiquitously expressed in the body and accumulates across cell types. It clearly affects the function of many cell types and yet it causes selective loss of particular neurons only. In fact, not even all somatic motor neurons are affected in ALS, but certain subpopulations show a strong resilience, including the oculomotor (cranial nerve 3, CN3), trochlear (CN4) and abducens (CN6) motor neurons, which innervate the extraocular muscles that control eye movement. This has been demonstrated both in patients (Caligari et al. 2013; Gizzi et al. 1992; Kubota et al. 2000), in mutant SOD1 ALS mice (Comley et al. 2016; Tjust et al. 2012; Valdez et al. 2013) and in inducible mutant TDP-43 mice (Spiller et al. 2016). Thus, this resilience is present across disease causations and species. Furthermore, visceral motor neurons that innervate smooth muscle appear relatively unaffected in ALS patients (Shimizu et al. 2011). Why particular neurons are more resilient than others to ALS and if this is entirely intrinsically programmed or also due to differences in non-cell autonomous factors in the environments surrounding motor neurons or in their muscle target innervation sites is still unclear. Through removing mutant SOD1 from particular cell types in transgenic mice, it has been demonstrated that mutant SOD1 within motor neurons themselves is important for driving initiation and early progression of disease (Boillee et al. 2006), while microglia and

astrocytes are important for later disease progression stages (Boillee et al. 2006; Yamanaka et al. 2008). Stem cell-derived oculomotor neurons also show a higher level of resistance to ALS-like toxicity *in vitro* indicating that part of their resilience comes from within (Allodi et al. 2018). Furthermore, we have demonstrated that oculomotor neurons express relatively high levels of different growth factors at baseline and in response to disease, as shown for spinal muscular atrophy (SMA), which can protect motor neurons and their synapses with muscle (Allodi et al. 2016; Nichterwitz et al. 2020). Oculomotor neurons also lack baseline expression of Mmp9, which particularly vulnerable spinal motor neurons express high levels of and which appears detrimental to motor neurons (Kaplan et al. 2013). Thus, it is likely the presence as well as lack of certain factors within neurons as baseline, in combination with their response to disease, that render a cell particularly resilient versus vulnerable to disease-stressors. Several studies have analyzed the transcriptional dysregulation that occurs within vulnerable spinal motor neurons in mutant SOD1 mouse models of ALS either prior to overt clinical symptoms (Perrin et al. 2005; Lobsiger et al. 2007), or at symptomatic stages (Perrin et al. 2005; Sun et al. 2015; Shadrach et al. 2021). These studies have shown that vulnerable neurons respond to disease by activating pathways of compensatory regeneration and response to DNA damage and cell injury.

No attempts have yet been made to comprehensively analyze how different types of vulnerable neurons respond to ALS over time in comparison to different types of resilient neurons, in order to disseminate the mechanisms by which particular neurons are vulnerable while others remain resilient. We reasoned that an analysis of how both vulnerable and resilient neurons respond to mutant SOD1 over time, as well as careful consideration of their baseline gene expression, could give further insight into disease mechanisms. We therefore set out to analyze different types of

resilient motor neurons, including somatic oculomotor (CN3) and trochlear (CN4) motor neurons, and visceral motor neurons of the dorsal motor vagus nerve (CN10) to delineate if all resilient neurons respond in a similar fashion to ALS. We compared two vulnerable motor neuron populations; lumbar spinal motor neurons and hypoglossal motor neurons (CN12), to also answer the question if vulnerable neurons share destructive pathways and attempts of compensatory responses or if the brainstem and spinal cord are going through distinct disease mechanisms, due to differences in target innervation and local environments. We used laser capture microdissection and Smart-seq2 RNA sequencing (LCM-seq) (Nichterwitz et al. 2016, 2018) to analyze the transcriptional dynamics across neuronal populations in ALS mice and wildtype littermates and confirmed DEGs using RNAscope. We also conducted machine learning and meta-analysis to identify strong disease predictors and a general vulnerability code across mutations and species.

Results

Transcriptional regulation of mutant SOD1 or disease associated genes do not underlie selective vulnerability among motor neurons

To retrieve temporal mechanistic insight into the selective vulnerability and resilience across motor neuron populations in ALS we isolated somatic motor neurons from CN3/4 (resilient), CN12 (vulnerable), and lumbar spinal cord (highly vulnerable) as well as visceral motor neurons from CN10 (resilient) of presymptomatic (P56) and onset of symptoms (P112) SOD1G93A mice and wild-type littermates using laser capture microdissection. We subjected the neurons to Smart-Seq2 polyA-based RNA sequencing (Fig. 1A, Supplemental Fig. S1). The majority of reads

across samples were uniquely mapped to the mouse genome, with a minority being multi mapped or non-mapped (Supplemental Fig. S2A). Only uniquely mapped reads were used for all further analyses. Pairwise Pearson correlation of the variance stabilized transformed (VST) expression values of all samples that passed the quality control (see *Methods)* shows a Pearson correlation of at least 0.95 within each cell type and did not show any separation of disease versus control within the respective cell types (Fig. 1B).

Hierarchical clustering of samples based on *Phox2a/b* and *Hox* gene expression shows a clustering based on cell type and position along the anterior-posterior body axis and confirms that CN3/4 neurons cluster based on *Phox2b* expression and lack of *Hox* gene expression, while spinal motor neurons clustered based on caudal *Hox* gene expression (*Hox6-11*) compared to CN10 and CN12 neurons, as expected (Supplemental Fig. S2B). Further marker gene expression analysis showed that all isolated motor neuron groups expressed *Chat* and *Islet1* (*Isl1*). CN12 and spinal motor neurons also expressed *Isl2* and *Mnx1*, while those markers were nearly absent in CN3/4 and visceral CN10 motor neurons (Supplemental Fig. S2C). All neuron groups expressed high levels of *peripherin* (*Prph*) and *neurofilament heavy chain* (*Nefh*), although CN10 had markedly lower levels of *Nefh* than the other groups (Supplemental Fig. S2C). Principal component analysis including all expressed genes showed that samples cluster based on cell type (Fig. 1c). The first principal component (PC1) clearly distinguished visceral (CN10) from the somatic (CN3/4, CN12 and spinal) motor neurons. The PC1 also to a lesser extent distinguished brainstem motor neurons from spinal motor neurons, while PC2 distinguished CN3/4 motor neurons from the other groups (Fig. 1C). Within the CN10 group, the PC2 axis also seemed to resolve age, but this was not the case for any other neuron group. ALS neurons were not clearly separated from

control neurons in this plot of all expressed genes (Fig. 1C), as a minority of genes were dysregulated in ALS compared to the overall gene expression across these neurons (Fig. 2A). The ALS mouse model is based on overexpression of human mutated *SOD1G93A* and the level of the resulting protein is important for disease development. Analysis of human and mouse *Sod1* levels in mutant SOD1G93A mice and wildtype littermates clearly shows that human *SOD1G93A* mRNA is highly expressed in the SOD1G93A mice across neuron types and that the endogenous mouse *Sod1* mRNA is expressed, albeit at lower level than mutant*SOD1G93A*, across all mice, wild-type and SOD1G93A, as expected (Fig. 1D). As the level of mutant *SOD1* is an important disease predictor we conducted a careful statistical analysis of any variation in expression level across cell types. We found only very minor variations in *SOD1G93A* levels across cell types, with CN10 and CN12 motor neurons showing the highest levels (Supplemental Fig. S2D). We conducted an ANOVA analysis comparing the two homologs of the *Sod1* gene, followed by a post-hoc Bonferroni-corrected t-test. The pairwise comparisons revealed significant but subtle differences between groups (Supplemental Fig.S2D). Collectively, these small differences in mRNA level of *SOD1G93A* do not appear to reflect the level of vulnerability of the respective cell types. Analysis of endogenous mouse *Sod1* mRNA confirmed that it was expressed at a lower level than human *SOD1G93A*, and demonstrated that there was only a very slight difference in levels across the cell types (Supplemental Fig. S2D). Thus, differential regulation of transcripts other than *SOD1G93A* and *Sod1* hold the key to understanding differential vulnerability in SOD-ALS. Analysis of mouse embryonic stem cell derived neurons have indicated that cranial motor neurons have higher proteasome activity to degrade misfolded proteins compared to spinal motor neurons and thus would be more resilient to proteostatic stress and misfolded SOD1 (An et al. 2019). To investigate

this matter *in vivo*, we analyzed GO terms related to proteasome, ubiquitination and proteolysis across samples of CN3/4 and spinal motor neurons in control and SOD1G93A mice, using Gene Set Variation Analysis (Hänzelmann S. et al. 2013, GSVA), but found no indication that CN3/4 motor neurons would have a better capacity to degrade proteins than spinal motor neurons in either control or SOD1G93A mice (Supplemental Fig. S2E). We subsequently analyzed the expression levels of other disease-causing or disease-modifying genes, including *Tdp43*, *Fus*, *C9orf72*, *Tbk1*, *Stmn2* and *Atxn2*, but found no variation across either cell types or disease states across cell types (Fig. 1D). In conclusion, different motor neuron types are clearly distinguished by their transcriptomic profiles and *SOD1G93A*-induced gene expression changes do not shift the identity of the motor neurons in any major way. We found that *SOD1G93A* is expressed at comparable levels across neuron types, and thus that differences in *SOD1G93A* expression levels do not reflect differences in susceptibility, nor did the expression level of other disease-causative or modifying genes. Therefore, to understand selective vulnerability, further analysis of genes and pathways dysregulated downstream of mutant *SOD1G93A* may reveal how cells respond differently and thus either succumb to or are shielded from disease.

Motor neuron subpopulations show unique spatial and temporal gene regulation in response to mutant SOD1 expression and resilient neurons appear rather unphased by disease

To retrieve mechanistic insight into the selective vulnerability and resilience of particular motor neuron groups in ALS, we analyzed differential gene expression induced by mutant *SOD1G93A* overexpression (Supplemental Table S1). Strikingly, the resilient groups, CN3/4 and CN10 regulated very few genes with disease compared to the vulnerable CN12 and spinal motor

neurons (Fig. 2A). At the presymptomatic P56 time point, CN3/4 showed 20 differentially expressed genes (DEGs) (15 upregulated and 5 downregulated), CN10 had 43 DEGs (32 up and 11 down), while CN12 had 169 DEGs (39 up and 130 down), and spinal motor neurons 146 DEGs (67 up and 79 down). At the symptomatic P112 stage CN3/4 showed 27 dysregulated genes (14 up and 13 down), CN10 had even fewer with 15 DEGs (6 up and 9 down), while CN12 had 110 DEGs (95 up and 15 down) and spinal motor neurons 234 DEGs (129 up and 105 down). Interestingly, CN12 gene regulation went from mainly suppression of gene expression at P56 to a majority of gene upregulation at P112 (Fig. 2A).

Analysis at the presymptomatic P56 stage, demonstrated that each motor neuron population had a unique response to mSOD1, with very little overlap in DEGs across cell types. Thus, neither vulnerable (CN12 or spinal) nor resilient (CN3/4 or CN10) motor neuron groups respond similarly to SOD1-ALS (Fig. 2B, Supplemental Fig. S3G). At the symptomatic P112 stage, most DEGs were still unique to the particular motor neuron group. Resilient CN3/4 and CN10 neurons only shared *mSOD1* as a DEG and thus clearly do not show a common induced resilience signature to ALS. However, the vulnerable CN12 and spinal motor neurons showed a larger overlap in DEGs in the symptomatic than at the presymptomatic stage, with an overlap of 34 genes, which represented 20% of the DEGs in CN12 and 14.5% of the spinal motor neuron DEGs (Fig. 2C).

Resilient ocular motor neurons upregulate a handful of genes with known neuroprotective properties in response to disease

To understand the spatially restricted temporal gene expression changes in disease we analyzed the DEGs within each motor nucleus across the P56 and P112 time points. Starting with the

resilient motor neurons, we identified 13 DEGs unique to CN3/4 at the presymptomatic stage which recovered to baseline at the symptom onset time, including *Gad1*, *Gabrq*, *Pcdh11x*, *Dcn*, *Scn3b*, *Slc17a6*, *Trim30e-ps1*, *Zfp982*, *Adcyap1* and *Gad2*. Among these, *Gad1* (glutamic acid decarboxylase), which converts glutamate to GABA was downregulated (Fig. 2D; Supplemental Fig. S3A), while *Gabrq* (gamma-aminobutyric acid type A receptor subunit theta) showed the opposite regulation with an upregulation at P56 (Fig. 2D, Supplemental Fig. S3B). Notably, *Gabrq* levels were pronouncedly higher in CN3/4 and CN10 than in vulnerable motor neuron groups (Supplemental Fig. S3B). The voltage-gated ion channel subunit, *Scn3b,* was downregulated in CN3/4 motor neurons only at the presymptomatic stage (Fig. 2D). The vesicular glutamate transporter *Slc17a6* was downregulated in presymptomatic CN3/4 and at both stages in CN12 motor neurons, while upregulated in symptomatic spinal motor neurons. The cadherin *Pcdh11x* was upregulated in CN3/4 at the presymptomatic stage while down in CN12 at the symptomatic P112 stage. The extracellular matrix (ECM) proteoglycan decorin (*Dcn*) was dramatically downregulated in CN3/4 motor neurons at the presymptomatic stage. Notably, *Dcn* levels were in general several fold higher in vulnerable neurons, particularly in CN12, in both health and disease, differences that were even more pronounced by the downregulation in CN3/4 motor neurons with early disease (Supplemental Fig. S3C). In CN3/4, seven of the 20 genes initially dysregulated were shared with the later time point where 27 genes were dysregulated. Of these commonly regulated genes *Ucn* (urocortin) (Pedersen et al. 2002), *Cck* (cholecystokinin) (Wozniak et al. 2021) and the ECM glycoprotein *Postn* (periostin) (Shimamura et al. 2012), have known neuroprotective properties. The levels of *Postn* were not only upregulated in CN3/4 motor neurons with disease (Fig. 2D), but were also several fold higher in CN3/4 motor neurons than all

other groups at baseline in control mice (Supplemental Fig. S3D). The distinct expression profiles of *Dcn* and *Postn* in vulnerable versus resilient motor neuron groups is compelling and warrants further investigation (Supplemental Fig. S3C,D). *Cartpt* (CART prepropeptide), *Dlk1* (delta like non-canonical notch ligand 1), *Scg2* (secretogranin II) (Fig. 2D) and *SOD1G93A* (not shown) were the remaining shared CN3/4 DEGs. The upregulation of *Dlk1* in CN3/4 motor neurons was shared with spinal motor neurons at P56, but not at P112. Interestingly, the level of *Dlk1* was significantly higher at baseline in both resilient motor neuron groups, CN3/4 and CN10, compared to vulnerable motor neurons, in control animals and in ALS (Fig. 2E). Twenty genes were uniquely regulated in CN3/4 motor neurons at the symptomatic stage, of which *Pgr15l, Tent5a, Fos and Npb* were upregulated while *Hbb-bt*, *Aldh1a3*, *Xaf1*, *Hbb-bs*, *Opn4*, *Chl1*, and *Gal* were downregulated. Notably, *Gal* and *Chl1* showed the opposite regulation in spinal motor neurons, where they were induced (Fig. 2D).

For CN10, 43 DEGs were identified at P56 and 15 DEGs at P112 with only three genes, *Gcg* (glucagon), *Npy* (neuropeptide Y) (Fig. 2F) and *SOD1G93A* (not shown), shared between the two disease stages. CN10 motor neurons regulated two genes, *Glra2* and *Slc34a2*, in a similar direction with vulnerable CN12 and spinal motor neurons. *Penk* (proenkephalin) on the other hand was downregulated in CN10 (at P112) and CN12 motor neurons (at P56), while upregulated in spinal motor neurons. *Uts2* on the other hand was upregulated in CN10 and CN12 populations while unchanged in spinal motor neurons with disease (Fig. 2F).

To elucidate any dysregulated cellular pathways with disease in the resilient populations of motor neurons, we conducted a comprehensive analysis employing three distinct generations of

pathway analysis methods: Overlap-Based Analysis (OVA) utilizing the EASE method, Per Gene Analysis (PGA) employing fGSEA, and Network Analysis (NA) conducted with the ANUBIX method. Each method offers unique insights into the respective biology, collectively contributing to a global understanding. No pathways were identified in CN3/4 motor neurons while in CN10 motor neurons four pathways were revealed at P112, and included 'adenylate cyclase activating G protein coupled receptor signaling pathway', 'ion transmembrane transport', 'regulation of membrane potential' and 'cell activation' (Supplemental Fig. S3H, Supplemental Table S2). In conclusion, based on the DEGs identified in resilient neuron groups it is evident that there is no common induced resilience signature across neuron types. Some of the DEGs identified, including *Ucn*, *Cck*, and *Postn*, point to neuroprotective pathways that may explain some of the tolerance towards mutant *SOD1G93A*. We also identified gene expression changes with as yet unclear roles, some of which at a first glance would appear to related to vulnerability e.g. *Dlk1* gene activation, which is a known marker of vulnerable spinal motor neurons (Müller et al. 2014), but that may be part of the neurons' innate function and/or resilience that does not necessarily put a cell at risk.

Analysis of basal gene expression differences highlights fundamental differences between CN3/4 and spinal motor neurons on top of which gene expression changes occur in ALS

The combination of few gene regulation changes occurring in resilient neurons, and no pathways identified in CN3/4 motor neurons in response to *mSOD1*, and the presence of high levels of *Dlk1,* which was previously described to be highly expressed in vulnerable fast fatigable (FF) motor neurons (Müller et al. 2014), prompted us to investigate the general basal gene expression in

resilient neurons. The underlying hypothesis was that the basal gene expression may hold the key to vulnerability, and explain why resilient neurons do not necessarily need to modulate their transcriptomes to a large extent in response to toxic gene expression. To investigate this matter, we first analyzed the number of expressed genes in CN3/4 versus spinal motor neurons and any commonality versus differences in basal gene expression patterns (Fig. 3A). The vast majority of transcripts were shared across neuron groups and times, as expected. CN3/4 motor neurons had 1218 uniquely detected transcripts shared across the two time points and an additional 782 transcripts unique to P56 and 733 genes unique to P112, which we reasoned could underlie their relative resilience to ALS (Fig. 3A). The transcripts uniquely enriched in either motor neuron population were visualized using Volcano plots (Fig. 3B,C, Supplemental Table S3). Spinal motor neurons showed, as expected, an enrichment of e.g. *Hox8-10* family members, *Mmp9*, *and Uts2* while CN3/4 motor neurons were enriched for *Phox2b*, *Eya1*, *Gabra1*, *Chrm1*, *Chrm3*, *Ucn*, *Tbx20, Postn, Irx2, Dlk1, Ctxn2, Col5a2, Chrna4, Krt25, Kcnd3, Glra2,* and *Cartpt* (Supplemental Fig. S4). The cell type specific marker enrichment identified is in line with previously published studies (Hedlund et al. 2010).

We subsequently analyzed GO terms enriched in control CN3/4 or control spinal motor neurons at P112, and found that 18 GO terms were enriched in CN3/4 neurons, while 260 were enriched in spinal motor neurons (Fig. 3D, Supplemental Table. S4). Here the ‘positive regulation of cation channel activity‘ (Fig. 3E), ‘synaptic signaling‘ (Fig. 3F) and ‘circadian rhythm‘ (Fig. 3G) appeared particularly interesting.

Our complementary analysis of the baseline expression in oculomotor and spinal motor neurons shows a large set of gene expression differences at baseline. We highlighted the fundamental

differences in ECM components between resilient and vulnerable neurons such as *Dcn* and *Postn,* which could have a major impact on neuronal structure, function, and connectivity. GO terms such as 'synaptic signaling', 'cation channel activity' and 'circadian rhythm' enriched in CN3/4 and pathways related to protein modification processes, metabolic processes and 'regulation of anatomical structure morphogenesis' enriched in spinal motor neurons highlight their fundamental differences on top of which gene expression shifts occur with disease, mainly in vulnerable neurons. Our analysis clearly points out that basal cell type-specific gene expression needs to be considered in addition to differential gene expression with disease. If you do not understand the baseline it is difficult to comprehend the biological significance of disease-linked gene expression changes.

Vulnerable motor neurons show unique regulation of injury response genes indicative of cellular stress, tissue remodeling and motor neuron subtype switching

We next set out to dissect gene expression patterns in relatively vulnerable motor neuron groups to investigate potential temporal and spatial overlap in gene regulation. Our initial analysis clearly demonstrates that gene dysregulation also in vulnerable neurons is tightly regulated both spatially and temporally. Thus, CN12 and spinal motor neurons show mainly transcriptional responses to *SOD1G93A* overexpression that are unique to the cell type and disease state (Fig. 4A). Thus, to fully understand disease it is necessary to carefully study each stage as it is difficult to predict the next step. Nonetheless, some of the shared responses across time may give deep insight into strong disease predictors and drivers.

For CN12, only seven out of 169 DEGs from P56 were shared with P112 and these were *Crh*,

Slc17a6, *Spp1*, *Marcks*, *Glra2*, *Gm3788* (Fig. 4A,B) and *mSod1* (not shown). Of these, *Slc17a6* and *Glra2* were also induced in spinal motor neurons with disease. *Slc17a6* was also induced in CN3/4 motor neurons at P56 and *Glra2* in CN10 motor neurons at P112 (Fig. 4B). Interestingly, symptomatic CN12 motor neurons showed a higher overlap in gene dysregulation with same age spinal motor neurons than with the P56 CN12 motor neurons, indicative of a shared response to disease across vulnerable populations at the later stage. To investigate this general vulnerability code, we analyzed the 34 DEGs shared across CN12 and SC at P112 (Fig. 4C). We identified a number of genes uniquely upregulated with disease across vulnerable motor neurons, but not in resilient neurons, including e.g. *Atf3*, *Sprr1a*, *Gal*, *Timp1*, *Cd44, Ngfr, Nupr1, Calb, Timp1, Ccl2, Ccl7 and Chl1* (Fig. 4C,D). For spinal motor neurons, 40 DEGs were shared across the time points, while 106 DEGs were unique to the presymptomatic P56 time point and 194 DEGs to the symptomatic P112 stage. Some of the 40 DEGs shared across time points were uniquely regulated in spinal motor neurons, including *Fgf21*, *Mt1, Sv2a, Chodl, Gpnmb, Mmp12, Syt7, Map1a, Atf4, Atp1a1* and *Ddit3* (Fig. 4D). *Chodl*, which is a known marker of FF motor neurons (Enjin et al. 2010) was consistently downregulated with disease in spinal motor neurons, while *Sv2a,* a marker of slow motor neurons, was upregulated (Fig. 4D). Some transcripts belonging to ongoing programmed neuronal death were only regulated in the later stages in spinal motor neurons and not seen in CN12 including *Adcyap1* and *Lcn2* (Fig. 4D, E).

To validate genes enriched in the vulnerable CN12 and/or spinal motor neurons in ALS we used RNAscope fluorescence *in situ* hybridization. A probe detecting *Vacht* (vesicular acetylcholine transporter) was used to identify lumbar spinal motor neurons. The upregulation of *Atf3* in vulnerable motor neurons (Fig. 4C) was confirmed by RNAscope at both the presymptomatic

stage (Fig. 4F; one-tail mean permutation test, *P*= < 0.00001 at P56) and at the symptomatic stage (Fig. 4G; one-tail mean permutation test, *P*= < 0.00001 at P112). Consistent with axonal sprouting, *Gap43* expression was also increased in both vulnerable populations at the symptomatic stage (Fig. 4C-E). *Sprr1a* was found upregulated in both vulnerable populations (Fig. 4C-E) and RNAscope on spinal cord sections confirmed the upregulation with disease at the symptomatic stage (Fig. 4H one-tail mean permutation test, *P*=< 0.00001). *Vgf* was induced in both vulnerable populations at P112 and this upregulation was confirmed by RNAscope in spinal motor neurons (Fig. 4I, one-tail mean permutation test, *P*=< 0.00001). We also confirmed the upregulation of *Timp1 at P112* (Fig. 4J, one-tail mean permutation test, *P*=< 0.00001), and *Pvalb*, which was upregulated in spinal motor neurons with disease (Fig. 4D) was also confirmed by RNAscope at the presymptomatic stage (Fig. 4K, one-tail mean permutation test, *P*=< 0.00001). *Fgf21* was confirmed upregulated at P56 (Fig. 4L, one-tail mean permutation test, *P*=< 0.00001) and P112 (Fig. 4M, one-tail mean permutation test, *P*=7 e-7). Thus, the *in situ* probe data were concordant with our RNA sequencing analysis, showing significant upregulation in SOD1G93A motor neurons. Our results so far reveal that there is a stress gene signature which is unique to vulnerable neurons, as resilient neurons do not show this upregulation. This signature is to a large extent shared across different vulnerable neurons, but the timing is distinct in their response to disease, with CN12 motor neurons showing a later onset of disease than spinal motor neurons. The lower overlap of DEGs in CN12 over time than spinal motor neurons is also likely a reflection of that this nucleus is not yet as far along in the disease process and thus general detrimental and regenerative processes are not activated at the presymptomatic stage, but just at the symptomatic stage, while in spinal motor neurons they are activated across the disease stages.

Pathway analysis indicates that vulnerable motor neuron subpopulations degenerate in a similar fashion but at distinct temporal paces

To elucidate the pathways activated in vulnerable motor neuron populations, we conducted a comprehensive enrichment analysis (EA) employing methods from three different categories: OVerlap Analysis (OVA), Per-Gene score Analysis (PGA) and Network Enrichment Analysis (NEA) methods, as for the resilient populations. We selected EASE (Hosack et al. 2003; OVA), fGSEA (Korotkevich et al. 2021; PGA) and ANUBIX (Castresana-Aguirre et al. 2020; NEA), being representative of their category. NEA methods stand out from other approaches due to their consideration of the interconnections among DEGs within the context of functional gene sets in a network framework (Buzzao et al. 2024). Pathways showing enrichment were identified using an FDR cutoff of 0.1. The pathways demonstrating the most robust enrichment, as indicated by consensus among all three methods, included e.g. 'regulation of neuronal death', 'inflammatory response', 'regulation of ERK cascade', 'MAPK cascade', 'regulation of cell adhesion', 'cell migration and synaptic signaling' (Fig. 5A, B). Further analysis demonstrated that 11 genes (8 of these DEGs were commonly regulated at P56 and P112 in spinal motor neurons), including *Atf3, Cd44, Gadd45a, Ngfr, Ccl2, Ccl7, Gal, Timp1, Nupr1, Serpinb1a* and *Chl1* drive six of the major affected pathways (Fig. 5C). To further investigate the commonality in gene network activation in CN12 and spinal motor neurons with disease, we analyzed the DEGs in the Funcoup mouse network based on all evidences described by the authors (Alexeyenko and Sonnhammer, 2009; https://funcoup.org/search/), which demonstrated the connectivity between *Gadd45*, *Ngfr*, *Atf3* and *Cd44* in the MAPK cascade and integration with *Chl1*, here designated to the 'Cell Neuron Death pathway, *Nupr1* in the 'Cell Death pathway' and *Serpin1b* in the 'Cell inflammation

pathway' (Fig. 5D). Moreover, using the DEGs related to common detrimental pathways in CN12 and SC at P112, we were able to see a clear temporal divergence along PC1 in PCA plot (Fig. 5E). In conclusion, our analysis shows that vulnerable neurons activate common pathways that are linked, but with clear temporal differences. This analysis solidifies our finding that CN12 motor neurons are not as far along in the disease process as spinal motor neurons in the symptomatic phase but progressing on a similar path.

Machine learning and meta-analysis across SOD1 mutations and models identify strong disease predictors

Next, we set out to reveal which of the 129 upregulated DEGs identified at symptom onset in spinal SOD1G93A motor neurons in our data set (Fig. 2A) would be the strongest predictors of disease in SOD1-ALS. Towards this purpose we used a single cell RNA sequencing data set, *Namboori et al. 2021,* where neurons were derived from induced pluripotent stem cells (iPSCs) harboring a *SOD1E100G* mutation or from corrected control iPSCs, at a time when the *SOD1E100G* motor neurons were starting to degenerate in cell culture (Fig. 6A). From this single cell RNA sequencing data set we selected motor neurons only, based on the co-expression of *SLC18A3* and *ISL1*. We used a random forest classifier (Breiman. 2001) as our machine learning approach to evaluate if our DEGs could classify these motor neurons into ALS or control (Fig. 6B). Our classification model achieved an average sensitivity of 71.8% (P=0.14), specificity of 73.1% (P=0.02), PPV of 50.5% (P=0.04), NPV of 88.2% (P=0.09), AUC of 83% (P=0.06), and accuracy of 72.8%. These classification performance measures were obtained using a rigorous nested Leave-One-Out (LOO) cross-validation approach. This method ensures that the evaluation of

performance is based on left-out samples rather than the training data.

These results demonstrate the effectiveness of our model in distinguishing between different cell genotypes using our upregulated DEGs at symptomatic age. We also identified the genes with the strongest importance for the classification which were *VGF*, *PENK*, *NTS (neurotensin)*, and *INA (internexin neuronal intermediate filament protein alpha)* (Fig. 6C). Other genes contributing to the classification, *SV2A* and *GAP43,* indicate motor neuron subtype switching and/or enrichment with disease and regeneration*,* respectively.

We next set out to evaluate if these disease-predicting DEGs were commonly regulated in spinal motor neurons in response to mutant *SOD1* across other published data sets. We reasoned that shared regulation across mouse cohorts, mutations, time points and even species would reveal strong disease regulators of great importance. Such a meta-analysis across models and SOD1 mutations has not been done previously and is clearly challenging, as there is a temporal regulation in gene expression with disease, as seen in our SOD1G93A data. Nonetheless, we did also identify overlap across time points particularly in the vulnerable neuron groups, speaking to the feasibility of such an analysis. We thus analyzed three additional published data sets that isolated RNA from spinal motor neurons from mice overexpressing human mutant *SOD1* and one where the RNA originated from human stem cell-derived neurons harboring a point mutation in *SOD1*; *i*) *Lobsiger et al. 2007* isolated motor neurons from lumbar spinal cord from onset of symptoms (15 weeks of age) in a fast progressing SOD1G37R mouse (reaches end-stage at 6 months) using LCM and Affymetrix Arrays; *ii*) *Sun et al. 2015* used TRAP to isolate mRNAs on ribosomes in motor neurons at symptom onset (8 months of age) in a slow progressing SOD1G37R mouse (reaches end-stage at 12 months) and *iii*) *Shadrach et al. 2021* used TRAP to

isolate RNA on ribosomes in motor neurons in symptomatic SOD1G93A (4 months of age) (Fig. 6D).

The four studies, including the present one, based on RNA isolated from motor neurons of mutant SOD1 mice shared 32 DEGs (Fig. 6E). Among these shared DEGs, we also found *Vgf*, *Gap43*, *Ina* and *Abca1* that all aided in the classification of human motor neurons into ALS or control. *Penk*, which was a strong disease predictor in the human data set (Fig. 6C) was only identified in our data set and that of *Lobsiger et al. 2007* (Fig. 6F). Only two DEGs were found across all five studies (mouse and human) and those were *Vgf and Tpd52l1* (Fig. 6F). While only upregulated DEGs were used to classify the human data set (Fig. 6C) *Tpd52l1* was consistently downregulated across all data sets, and thus may also predict disease (Fig 6F). *Vgf* is a DEG that was also upregulated in CN12 with disease, but remained unaffected in resilient neuron groups (Fig. 3B). While *Vgf* was a DEG across all studies, it was upregulated in the current study (*Mei et al.),* in *Lobsiger et al. 2007*, in *Shadrach et al. 2021*, and in *Sun et al. 2015*, it was down-regulated in *Namboori et al. 2021*. (Fig. 6G). A previous study has shown that the VGF protein level in cerebrospinal fluid (CSF) was decreased in ALS patients by approximately 40% and VGF immunoreactivity in *post mortem* spinal motor neurons was also decreased (Zhao et al. 2008). As we saw the opposite pattern on the mRNA level with disease across the mouse studies, we decided to analyze the VGF protein level using immunofluorescence staining of SOD1G93A spinal cords and wild-type littermates (N=5/genotype). The signal was quantified in an automated fashion and demonstrated that VGF protein was indeed decreased in ALS motor neurons with disease (Welch's two-sample t-test P= <2.2e-16, Fig. 6H), consistent with previous findings (Zhao

et al. 2008). Thus, the mRNA increase of *Vgf* may be a compensatory event due to the loss of VGF protein. In conclusion, we have identified strong ALS disease markers and predictors based on DEGs identified in our study and across multiple published data sets, which predict disease across SOD1 mutations and species.

Discussion

We conducted RNA sequencing analysis of motor neurons displaying differential vulnerability to degeneration in ALS to unveil mechanisms of protection and degeneration at distinct disease stages. Our analysis did not reveal any major differences in human mutant *SOD1* or mouse *Sod1* mRNA levels across neuron types that could explain their differences in vulnerability. Neither was there any significant differences in other disease-associated gene expression including *Tdp43*, *Fus*, *C9orf72*, *Tbk1*, *Stmn2*, or *Atxn2* or pathways related to protein processing. Thus, the cause for differences in susceptibility to ALS lies elsewhere and analysis of responses induced by mutant *SOD1* may give mechanistic insight into both why certain neurons succumb to disease and why others persist.

Analysis of differential gene expression consistently showed that resilient motor neurons display a very mild transcriptomic response to disease while vulnerable motor neurons showed large transcriptional dysregulation already early in disease. This is different from the gene dysregulation we previously noted in spinal muscular atrophy (SMA), where resilient CN3/4 motor neurons elicited a large early gene response compared to vulnerable spinal and facial (CN7) motor neurons which only responded to the loss of the *Smn1* gene (survival motor neuron 1) later in disease. In the SMA mouse model we saw an induction of a large set of intuitively

protective genes in CN3/4 motor neurons in combination with a disease module that was shared across vulnerable and resilient neurons (Nichterwitz et al. 2020). This indicates that CN3/4 motor neurons are not as severely affected by misfolded mutant SOD1 as they are to loss of SMN and consequent disruption of RNA splicing. It is also consistent with that CN3/4 motor neurons show less aggregation of misfolded SOD1 compared to vulnerable spinal motor neurons or CN12 motor neurons, as demonstrated in the SOD1G93A mouse (An et al. 2019) and in SOD1G85R mice (Thomas et al. 2018.)

Even though SOD-ALS resilient neurons, both CN3/4 and CN10, regulated very few genes and thus did not seem in need to modulate gene expression in a major way to survive, the gene expression changes that did occur may hold a partial key to their resilience. Seven genes were dysregulated across disease stages in CN3/4 motor neurons, including *Ucn* (urocortin), *Postn* (periostin), *Cartpt* (CART prepropeptide), *Cck* (cholecystokinin), *Dlk1* (delta like non-canonical notch ligand 1), *Scg2* (secretogranin II) and *mSod1*. Of these DEGs *Ucn* has been shown to protect hippocampal neurons from oxidative stress and excitotoxic glutamate insult at picomolar levels (Pedersen et al. 2002). *Cck* has been shown to protect Purkinje cells from toxicity in ataxias (SCA mice) (Wozniak et al. 2021). *Postn* (periostin) levels were not only upregulated in CN3/4 motor neurons with disease, but were also several fold higher in CN3/4 motor neurons than all other groups at baseline in control mice. POSTN is an ECM glycoprotein that can stimulate neurite outgrowth (Shih et al. 2014) and be neuroprotective as shown in cortical neurons (Shimamura et al. 2012). The distinct expression profiles of *Dcn* and *Postn* in vulnerable versus resilient motor neuron groups is very compelling and warrants further investigation as these may influence both their physiology, their connectivity, communication with neighboring cells and response to

disease. Dlk1 (Notch ligand Delta-like homolog 1) was initially described as highly enriched in fast motor neurons and required and sufficient to promote a fast biophysical signature, where loss shifted motor neurons to a slow identity (Müller et al. 2014). This may indicate that Dlk1 could be used to modulate resilience in ALS, as slow motor neurons, which have much less Dlk1, persist longer than fast, but this remains to be investigated. However, we found that resilient CN3/4 and CN10 motor neurons express higher levels of *Dlk1* at baseline than all other motor neuron groups, and this expression was further induced with disease. Furthermore, both VTA and SNc dopamine neurons, of which SNc neurons are highly vulnerable to degeneration in Parkinson's disease, express high levels of *Dlk1* (https://mouse.brain-map.org/experiment/show/71587885). Future studies would benefit from analyzing what impact knockdown of *Dlk1* has on the integrity across ALS-vulnerable, and resilient neuron types which may rely on DLK1 for their function.

To better comprehend why CN3/4 motor neurons do not need to regulate a large number of genes to cope with disease, we conducted a complementary analysis of the baseline expression differences between CN3/4 and spinal cord motor neurons in control across time. This analysis highlights the fundamental differences on top of which gene expression changes occur with disease and also shows that to understand disease we need to investigate the underlying baseline. Here it became evident that while CN3/4 motor neurons expressed a somewhat larger number of genes, albeit many of those at low levels, the number of GO terms uniquely enriched in each nucleus was overwhelmingly larger in spinal motor neurons with 260 enriched GO terms compared to only 18 in CN3/4 motor neurons. Thus, oculomotor neurons may require fewer processes for their function than spinal motor neurons do, maybe due to their lower metabolic

demand and output onto many fewer muscle fibers.

Several of the genes regulated in vulnerable spinal motor neurons were previously shown to be upregulated in spinal motor neurons in mutant SOD1 mice (Lobsiger et al. 2007). Our study reveals that this gene signature is unique to vulnerable neurons, as resilient neurons do not show this upregulation and that it is shared across SOD1 mutations, and is clearly part of a stress response to disease. Vulnerable spinal motor neurons showed an upregulation of a number of genes important for and indicative of nerve regeneration, including *Gap43*, *Cd44*, *Chl1*, *Adycyap1*, that mirror the events that take place during certain phases of disease, where slow motor neurons regenerate and compensate for the loss of FF motor neurons (Frey et al. 2000; Fischer et al. 2004; Schaefer et al. 2005), resulting in a temporary stabilization of NMJ loss versus gain (Comley et al. 2016). Reduction of Cd44 has been shown to reduce axon initiation of retinal ganglion cells (Ries et al. 2007) and thus an increase as seen in the SOD1G93A motor neurons is expected to promote axon growth. The upregulation of *Atf3* in spinal motor neurons across time points and at the late time point in CN12 is also likely protective as it has been shown that overexpression of *Atf3* promotes motor neuron sprouting and survival as well as retained innervation of muscle in ALS mice (Seijffers et al. 2014). *Sprr1a* is also an axon regeneration-induced gene (Starkey et al. 2009) and anticipated to be a protective response as is the upregulation of *Pvalb*, which can protect neurons from excitotoxicity (Van Den Bosch et al. 2002). The downregulation of *Chodl*, an FF motor neuron marker in spinal motor neurons and upregulation of the S motor neuron marker Sv2a, could be indicative of a difference in the proportions of FF versus S motor neurons that are isolated over time with disease as FF motor

neurons are highly vulnerable to ER stress (Saxena et al. 2009) and lost early while S motor neurons remain and compensate by sprouting in the mSOD1 mouse (Pun et al. 2006). Thus, some of the unique regulation at the late stage of disease in spinal motor neurons may represent a difference in motor neuron subtypes incorporated (during the LCM isolation) rather than an induction of gene regulation. Future single cell or nuclei RNA sequencing studies may be able to resolve this issue if a temporal analysis is conducted.

The upregulation of *Syt7* in only spinal motor neurons is indicative of their need to reorganize presynaptic function, while the upregulation of *Lcn2* on the other hand appears to be a way to negatively control regeneration (Gu et al. 2024), perhaps a way to block motor neurons that are already handling too much stress- and DNA damage responses from regenerating, but this remains to be further investigated. Hypoglossal motor neurons also showed an induction of regenerative genes including *Gap43*, *Cd44* and *Chl1*. Some transcripts belonging to a disease response were only regulated in the later stages in spinal motor neurons and not seen in CN12, including *Adcyap1*, which is also linked to nerve regeneration (Baskozos et al. 2020). This may both be a reflection of that CN12 motor neurons are not as far along into the disease process as spinal motor neurons and that they have cell-type specificity in their response to disease. To fully resolve this question, neurons isolated from later stage diseased animals would need to be studied.

The consequences of the upregulation of *Ccl2* and *Ccl7* in vulnerable neurons, a response that is shared across ALS and nerve injury, as shown by Shadrach and colleagues (Shadrach et al. 2021), is not completely clear. Literature indicates that it could lead to enhanced neuronal excitability

(Zhou et al. 2011), and glial activation in the CNS, (Joly-Amado et al. 2020), which may be detrimental.

Fgf21 was robustly and specifically upregulated in spinal motor neurons across disease stages. FGF21 is involved in regulating carbohydrate and lipid metabolism and maintaining energy homeostasis, and it can protect cells from apoptosis. FGF21 is induced by ER stress, mitochondrial dysfunction and starvation. We and others have recently shown that ALS motor neurons display mitochondrial dysfunction early on (Hor et al. 2021; Mehta et al. 2021; Schweingruber et al. 2023), and thus *Fgf21* upregulation may be a consequence thereof. The upregulation of *Fgf21* in the vulnerable spinal motor neurons may also be indicative of the increased energy demand on these cells during disease. There is a clear correlation between increased serum levels of FGF21 and metabolic diseases conditions such as diabetes, mitochondrial diseases, obesity and aging, all of which have muscle loss as a common factor (reviewed in Tezze et al. 2019).

The upregulation of *Nupr1* (nuclear protein 1) in CN12 and spinal motor neurons is indicative of the stress these cells are experiencing, as this transcriptional regulator is involved in ER stress, oxidative stress response, DNA repair, autophagy, apoptosis and chromatin remodeling (*reviewed in* Liu et al. 2022). In spinal motor neurons it was regulated already at the presymptomatic stage while in CN12 neurons only at the symptomatic stage, demonstrating that these cell types follow a similar path at distinct paces, where spinal motor neurons have taken the lead to destruction. Furthermore, *Ddit3* (DNA-damage inducible transcripts 3) (CHOP) is an ER-stress apoptotic mediator which was upregulated in spinal motor neurons alone, which is another clear indication that these cells are further along a degenerative pathway than CN12

motor neurons. *Ddit3* mRNA was previously shown to be upregulated in spinal motor neurons in 90-120 day-old SOD1G93A mice (Perrin. et al. 2005), and at 15 weeks in the SOD1G37R mouse (Lobsiger et al. 2007) as well as the protein level in end-stage sporadic ALS patient motor neurons (Ito et al. 2009). The earlier detection in our data, already at P56, may be a reflection of the sensitivity of Smart-seq2 and is in concordance with the very early ER stress response of fast fatigable (FF) motor neurons, seen prior to any visible denervation (Saxena et al. 2009). The upregulation of *Mt1* in spinal motor neurons across time points indicate a response to block apoptosis (Zhang et al. 2013), and the specific upregulation of *Gpnmb* appears to be part of an inductive protective response, as this molecule has been shown to be protective in ALS and upregulated in sera from sporadic ALS patients (Tanaka et al. 2012).

Extracellular matrix remodelling is regulated by the activity of MMPs (matrix metalloproteinases) which are tightly regulated by tissue inhibitors (Timps). The upregulation of *Timp1* (metallopeptidase inhibitor 1) across time points in spinal motor neurons and in the symptomatic stage in CN12 motor neurons, indicate again that these two vulnerable motor neuron groups follow similar paths of response and destruction albeit on a different time scale. *Mmp12* was also upregulated in spinal motor neurons indicating ongoing ECM remodeling.

In general, we note that many gene regulations in vulnerable neurons indicate compensatory events to handle the toxicity of mutant SOD1. The majority of these appear to be beneficial, such as nerve regeneration programs, but clearly insufficient over time. We also note that different vulnerable neuron populations share some responses to ALS but that their timing is distinct, likely

due to differences in their temporal involvement in the disease model. This was clearly demonstrated in our Enrichment analysis (EA), where we evaluated the three most represented methods of different categories of EA (Buzzao et al. 2024) within GOBP. Commonalities pathways in “regulation of neuronal death”, “inflammatory response”, “regulation of ERK cascade”, “MAPK cascade”, “regulation of cell adhesion”, “cell migration” and “synaptic signaling” exhibited consistent enrichment across all methodologies in vulnerable motor neurons.

To further understand the strength in the identified vulnerability signature in both defining and predicting disease across data sets, we compared our RNA sequencing data to that of five other published studies. We found that *Vgf* and *Tpd52l1* were identified as DEGs across all data sets independent of disease status, presymptomatic, or symptomatic, and across SOD1G37R, SOD1G93A and SOD1E100G mutations. A 4.8 kDa VGF peptide was previously described as a potential biomarker of ALS as the levels were decreased in CSF of ALS patients and could distinguish them from controls (Pasinetti et al. 2006). VGF is involved in energy expenditure so decreased levels have been hypothesized to contribute to the hypermetabolic state seen in ALS. A follow-up study confirmed VGF as a biomarker for ALS using ELISA which indicated that VGF CSF levels may be correlated with muscle weakness in ALS patients (Zhao et al. 2008). While *Vgf* mRNA was upregulated in our data and all other mouse transcriptome studies analyzed, we found VGF protein to be decreased in ALS motor neurons with disease in ALS mice, consistent with previous studies. The inverse correlation between RNA and protein levels indicates that there is a compensatory mechanism at play to increase VGF either by increased transcription, or by changes in RNA stability.

In conclusion we demonstrate that resilient motor neurons only regulate a few genes in response to mutant *SOD1* likely as their baseline gene expression renders them resilient against this specific insult. The few genes that were upregulated in CN3/4 have known protective properties, including *Ucn*, *Cck* and *Postn* and may confer resilience. We also reveal that different vulnerable motor neuron populations share pathway activation, including 'regulation of neuronal death', 'inflammatory response', 'ERK and MAPK cascades', 'cell adhesion', 'cell migration' and 'synaptic signaling', which indicate that cell death occurs through similar mechanisms across vulnerable motor neurons, but temporally separated. Machine learning and meta-analysis across mutant SOD1 data sets and disease time points reveal a shared transcriptional vulnerability disease code and identify *VGF*, *PENK, INA*, *GAP4*3 and *TPD52l1* as strong disease-predictors across SOD1 mutations and species. These genes thus have potential as future biomarkers of disease and may aid in diagnosis and prognostics.

Methods

Ethics statement and animal model

All procedures involving animals were approved by the Swedish ethics council, a local animal ethics committee (Paris CE5, France) and carried out according to the Code of Ethics of the World Medical Association (Declaration of Helsinki). Animals were housed with a 12-hour dark/light cycle under standard conditions and had access to food and water *ad libitum*. Adult SOD1G93A mice on a C57Bl/6J background (B6.Cg-Tg(SOD1-G93A)1Gur/J, The Jackson Laboratories strain # 004435, males) were used as a model of ALS and non-transgenic littermates served as a control. All animals were anesthetized with a lethal dose of avertin (2,2,2- Tribromoethanol in 2-Methylbutanol, Sigma-Aldrich) prior to either decapitation or intracardial perfusion with phosphate buffered saline (PBS) followed by 4% paraformaldehyde (PFA) in PBS. Animals used for RNA sequencing originated from a cohort of animals in Sweden, all animals used for RNAscope derived from a cohort located in France.

Tissue processing and laser capture microdissection for transcriptomics

For the dissection and processing of tissues for LCM-seq, all equipment, surfaces, and tools were carefully cleaned with RNaseZap (Ambion/Life Technologies) and wiped with distilled water or 70% ethanol. Brain and spinal cord tissues were dissected from 56 and 112-day old males and immediately snap-frozen in 2-Methylbutane (Sigma-Aldrich) on dry ice, followed by storage at -80°C until further processing. After equilibration of the tissue to approximately -20°C, 12 µm coronal sections were prepared on a cryostat and placed onto PEN membrane glass slides (Zeiss). Spinal cords were embedded in OCT (Dako) prior to sectioning. Slides were stored at -80°C until

further processing. For laser capture microdissection (LCM), cells were visualized with a quick histological staining (Histogene, Arcturus/Life Technologies) followed by dehydration in a series of ethanol solutions (75%, 95%, and 99.7%, Solveco). LCM was performed on a Leica LMD 7000 system at 40x magnification and cutting outlines were drawn in close proximity to individual cells to minimize contamination by surrounding tissue. Per sample, approximately 100-200 cells with an area of >200 μm^2 (150 μm^2 for vagus motor neurons) and a visible nucleus with nucleolus were collected into the dry cap of a PCR tube (Biozym Scientific). After the addition of 5 µl lysis buffer (0.2% Triton X-100, with 2 U/µl recombinant RNase inhibitor, Clontech) to the cap, samples were mixed by pipetting up and down, spun down in a tabletop centrifuge and snap frozen on dry ice. Samples were stored at -80°C until library preparation.

Tissue processing for immunohistochemistry and RNAscope

Male and female mice were transcardially perfused with PBS followed by 4% PFA in PBS at P56 or P112. The lumbar region of the spinal cord was dissected, post-fixed for 3 h in 4% PFA at 4°C, and cryoprotected in PBS 30% sucrose for 48 h at 4°C. Spinal cord sections 30 µm thick were prepared using Thermo Scientific Sliding Microtome Microm HM 430.

RNAscope fluorescent *in situ* hybridization

As previously described (Leboeuf et al. 2023), sections were washed in PBS, incubated with RNAscope hydrogen peroxide solution from Advanced Cell Diagnostics (ACD) for 10 min at room temperature (RT), rinsed in PBS 0.1% Tween 20 at RT, collected on Super Frost plus microscope slides (Thermo Scientific), dried at RT for 1 h, rapidly immersed in ultrapure water, dried again at

RT for 1 h, heated for 1 h at 60°C, and dried at RT overnight. The next day, sections were immersed in ultrapure water, rapidly dehydrated with 100% ethanol, incubated at 100°C for 15 min in RNAscope 1X Target Retrieval Reagent (ACD), washed in ultrapure water, and dehydrated with 100% ethanol for 3 min at RT. Protease treatment was carried out using RNAscope Protease Plus solution (ACD) for 30 min at 40°C in a HybEZ oven (ACD). Sections were then washed in PBS before *in situ* hybridization using the RNAscope Multiplex Fluorescent V2 Assay (ACD). Probes were hybridized for 2 h at 40°C in a HybEZ oven (ACD), followed by incubation with signal amplification reagents according to the manufacturer's instructions. Probes were purchased from ACD: Mm-Slc18a3-C3 (catalog #448771-C3), Mm-Atf3-C2 (catalog #426891-C2), Mm-Sprr1a-C2 (catalog #426871-C2), Mm-Timp1-C1 (catalog #316841), Mm-Fgf21-C1 (catalog #460931), Mm-Pvalb-C4 (catalog #421931-C4), Mm-Vgf-C2 (catalog #517421-C2). The hybridized probes' signals were visualized and captured on a Zeiss LSM800 Airyscan confocal microscope with a 20× objective.

Spinal cord immunohistochemistry

Sections were quickly washed in PBS and incubated in 5% donkey serum, 0.3% triton-X 100 in PBS for 1 h at RT. Sections were then incubated overnight with primary antibodies diluted in 5% donkey serum, 0.3% triton-X 100 in PBS at 4°C, washed and further incubated with secondary antibodies for 1 h at RT. Primary antibodies included rabbit anti-VGF (LSBio LS-C352987, 1:100), goat anti-VAChT (Millipore ABN100, 1:500). Controls without primary antibodies were included. Images were acquired on a Zeiss LSM800 Airyscan confocal microscope with a 20× objective.

Image analysis

All mouse SC sections at P56 and P112 were imaged at 20X on a confocal microscope (Axio-Observer Z1/7). Scans were acquired at high resolution (4084 × 4084) with a z-step size of 0.58 µm. Maximum intensity images were firstly generated from merged z-focal planes using CellProfiler (Stirling et al. 2021, v.4.2.5). Intensity of RNAscope probes and immunohistochemistry images were quantified within the masked VAChT/ChAT+ cells using an automated CellProfiler pipeline. Each cellular specimen exhibiting an intensity of VAChT/ChAT+ signal higher than 0.01 was subjected to quantification. Subsequently, the normality of the data distribution was assessed. Statistical testing was performed utilizing one tailed mean permutation test with a permutation count of 100,000 and the nominal p-value is presented.

Reagents

For RNA-seq experiments, all reagents were of molecular biology/PCR grade if available. Only nuclease-free water (H_2O, LifeTechnologies) and tubes that were certified nuclease-free were used. All workbenches and equipment were cleaned with RNaseZAP (Ambion/Life Technologies) and additionally with DNAoff (Takara) for library preparations.

cDNA and sequencing library preparation

For library preparation, a modified version of the Smart-seq2 protocol (Picelli et al. 2014b, 2013) was used, which is described in detail in (Nichterwitz et al. 2018, 2016). All heat incubation steps and PCR cycles were carried out in a BioRad T100 Thermal Cycler. After reverse transcription, cDNA was amplified for 18 cycles, followed by purification with magnetic beads (GE Healthcare).

cDNA concentration and library quality were assessed using an Agilent 2100 Bioanalyzer (High Sensitivity DNA kit). Three samples with poor cDNA library quality were excluded from analysis (low cDNA concentration with no apparent peaks). One ng of cDNA (as determined with the Bioanalyzer, 100-9000 bp range) was used as input for the tagmentation reaction, which was carried out with 0.4 -1 µl of in house Tn5 (Picelli et al. 2014a). Ligation of sequencing indices (Nextera XT Sequencing Index Kit, Illumina) and 10 cycles of enrichment PCR were performed with Kapa HiFi polymerase. Final sequencing libraries were purified with magnetic beads and the concentration of each sample was determined on a Qubit fluorometer (ThermoFisher) using the dsDNA high sensitivity kit (LifeTechnologies). Samples were pooled and sequenced on the Illumina Hiseq 2500 seq platform.

RNA-seq analysis

The RNA-seq reads were mapped simultaneously to the mm39 mouse genome assembly and the genomic sequence of human SOD1 from the hg38 assembly using STAR (version 2.7.0e) (Dobin et al. 2013). We used uniquely mapped reads (69.7 ± 0.50% mean± SEM) (Supplemental Fig.S2D) for further analyses. Expression levels were determined using the rpkmforgenes.py software (http://sandberg.cmb.ki.se/rnaseq) with the Ensembl gene annotation. Samples included in the analysis had more than 15,000 detected genes (16408 ± 75 genes; *mean ± SEM*). Genes with at least one count in at least five samples were retained in the dataset. Differential expression analysis was performed using the R-package EdgeR (v3.36.0 Robinson et. al. 2010). A quasi-likelihood negative binomial generalized log-linear model was fitted to the data before empirical Bayes moderated t-statistics were calculated and multiple testing correction (Benjamin-

Hochberg) was performed. The threshold false discovery rate (FDR) < 0.05 was used to determine if a transcript was differentially expressed. For Supplemental fig. S2C, an ANOVA was implemented with one factor (Cell type_ Age) which was significant. We then ran post-hoc analyses to assess pairwise differences and t-test results with Bonferroni correction are reported.

Enrichment analysis

The Enrichment analysis methods used in this study belong to three different categories: OVerlap Analysis (OVA), Per-Gene score Analysis (PGA) and Network Enrichment Analysis (NEA) methods. As methods, we selected EASE (OVA), fGSEA (Korotkevich et al. 2021, PGA) and ANUBIX (Castresana-Aguirre M et al. 2020, NEA). All methods were run in an R (v4.1.1) version-controlled conda environment using the original packages, EASE and ANUBIX were run as in a recent benchmarking study for gene set EA methods (Buzzao et al. 2024). EASE and ANUBIX were performed on differentially expressed genes with an FDR below 0.1. FunCoup (Alexeyenko and Sonnhammer, 2009; https://funcoup.org/search/) was used as a network with a default confidence link score of 0.8 and 5 max neighbors in order to obtain the network of links that exist between vulnerable genes within detrimental pathways.

Gene Set Variation Analysis (Hänzelmann S. et al. 2013, GSVA, v.1.50.1) was performed to assess pathway enrichment. GSVA was used on log-transformed counts per million (CPM) values calculated from the expression data. A GSVA parameter object was then created with the 'gsvaParam' function, using log-transformed CPM values, list of GO-BP gene sets (MSigDB), kcdf=Gaussian and maxDiff=False. The 'gsva' function was subsequently applied to compute enrichment scores for each sample.

Random forest classifier

The data preprocessing of Namboori et al 2021, included normalization using DESeq2 (v.1.34.0, Love et al., 2014), followed by scaling of the gene expression data using the scale() function. The scaling was applied to the genes, involving centering each gene's expression values around its mean and scaling to unit standard deviation (SD). To address class imbalance within our dataset, we employed the downSample() function from the ROSE package in R. The random forest model is trained using nested cross-validation to ensure robust evaluation. In the outer loop of the nested cross-validation the dataset is repeatedly split into training and test sets. Within each iteration of the outer loop, an inner cross-validation is performed using the train() function from the caret package, with the target variable being the factorized disease status (ALS or Control) and the predictors being the Namboori et al. 2021, gene expression data.

For the inner cross-validation, the train() function applies a specified resampling method (such as k-fold cross-validation) to the training data. It systematically partitions the training data into subsets, trains the model on a subset, and evaluates it on the remaining subset. This process helps optimize model hyperparameters and provides an unbiased estimate of model performance within each iteration of the outer cross-validation loop. During the outer cross-validation loop, the performance of the random forest model is assessed using metrics such as: accuracy, sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the ROC curve (AUC). These metrics are computed only for the left-out samples of the outer cross-validation test sets, ensuring an unbiased estimate of the model's performance. The importance of each gene in predicting disease status is assessed using

permutation importance, providing insights into the contribution of individual genes to the classification task.

Use of published datasets

To evaluate the purity of our LCM collected samples, we compared our data to a previously published dataset obtained from GEO, accession number GSE52564 (Zhang et al. 2014), which was processed as described for our samples. For the comparison of our data with other spinal motor neuron microarray and RNA-seq data, the previously published data were preprocessed as described below. The microarray data (Lobsiger et al. 2007) were analyzed after rma() normalization and quality control for both moe430A and B arrays. Shadrach et al. 2021 and Sun et al. 2015 were mapped to the mm10 and mm39 assemblies, respectively, using STAR (version 2.7.0e). Genes with a minimum count of 4 and 3, respectively, within all samples were retained. Namboori et al 2021, was mapped to the hg38 assembly and processed retaining all cells with co-expression of *Slc18a3* and *Isl1* of at least 5 counts and nUMI >= 1000, nGene >= 300 and mitoRatio<= 0.5). Differential analysis was carried out using Wald test in DESeq2 and parameters set as follows: sfType = "poscounts", minReplicatesForReplace = Inf, useT = TRUE and minmu = 1e-06.

Data access

All LCM-seq RNA-seq data generated in this study have been deposited at the Gene Expression Omnibus (GEO) of the National Center for Biotechnology Information, with the accession number GSE244538.

Acknowledgements

We would like to thank Chris Molenaar, facility manager at IFSU for providing excellent technical expertise on confocal imaging. We would like to thank Anna Klemm for her great assistance in the RNAscope images analysis. We would also like to thank Dr Rita Almeida (SU) for her insightful comments and suggestions regarding statistical analysis of RNAscope data. We would like to thank Mattias Karlén for his outstanding work on creating schematics in Figures 1 and 5. And we would also like to thank all members of the Hedlund and the Bartosovic laboratories for fruitful discussions and helpful suggestions regarding lab meetings. Confocal microscopy was performed at the Imaging Facility at Stockholm University (IFSU). The computations and data handling were enabled by resources provided by the National Academic Infrastructure for Supercomputing in Sweden (NAISS) and the Swedish National Infrastructure for Computing (SNIC) at UPPMAX. This work was supported by grants from the Swedish Research Council (2016-02112 and 2020-01049) to E.H., European Union Joint Programme for Neurodegenerative Disease (JPND) (529-2014-7500) to EH, from the Department of Biochemistry and Biophysics, Stockholm University to E.H., from Åhlén's Foundation (Åhlén-stiftelsen) to E.H., from Ulla-Carin Lindquists foundation for ALS research (Ulla-Carin Lindquists stiftelse för ALS forskning) to E.H., and Birgit Backmarks Donation to ALS research at Karolinska Institute (Birgit Backmarks Donation till ALS forskning vid Karolinska Institutet) to E.H. I.L. has been financed by a MESR ('French ministry for high education and research') PhD fellowship. Project related funding for C.S.L. comes from the french national ALS association ('Association pour la recherche sur la SLA - ARSLA') and additional ALS associations ('Aide à la Recherche des Maladies du Cerveau - ARMC', and 'SLA Fondation Recherche - SLAFR').

C.S.L. thanks the staff of the animal housing facility UMS28/CEF (Paris, France) and of the iGenSeq platform for genotyping (ICM, Paris, France; which received funding from the program 'Investissements d'avenir' ANR-10-IAIHU-06).

Author contributions

E.H. conceived the study. I.M., S.N., M.L., D.L and E.H. designed experiments. E.H., D.R. and C.L. supervised the project. I.M., S.N., M.L., J.N., I.L., and C.L., acquired data, specifically S.N., and J.N., collected tissues for LCM and isolated neurons using LCM. S.N. prepared cDNA libraries for sequencing; I.M., and J.N. mapped sequencing data. I.M. conducted all bioinformatics analysis of RNA sequencing data; C.L. and I.L prepared tissues for RNA scope; M.L. conducted RNA scope; M.L. and I.M. conducted RNA scope imaging and analysis. M.L. and I.M. conducted immunofluorescent analysis. I.M., S.N., M.L., J.N., C.L. and E.H. analyzed data; I.M., M.L., S.N., E.H. prepared all Figures for paper. E.H. wrote the initial manuscript with the help of I.M., M.L. and C.L. All authors edited and gave critical input on the manuscript.

Disclosure declaration

The authors declare no conflicts of interest.

References

Alexeyenko A, Sonnhammer ELL. 2009. Global networks of functional coupling in eukaryotes from comprehensive data integration. *Genome Res* **19**: 1107-1116. doi: 10.1101/gr.087528.108

An D, Fujiki R, Iannitelli DE, Smerdon JW, Maity S, Rose MF, Gelber A, Wanaselja EK, Yagudayeva I, Lee JY, Vogel C, Wichterle H, Engle EC, Mazzoni O. 2019. Stem cell-derived cranial and spinal motor neurons reveal proteostatic differences between ALS resistant and sensitive motor neurons. *Elife* **8**:e44423. doi: 10.7554/eLife.44423

Baskozos G, Sandy-Hindmarch O, Clark AJ, Windsor K, Karlsson P, Weir GA, McDermott LA, Burchall J, Wiberg A, Furniss D, Bennett DLH, Schmid AB. 2020. Molecular and cellular correlates of human nerve regeneration: ADCYAP1/PACAP enhance nerve outgrowth. *Brain* **143**: 2009-2026. doi: 10.1093/brain/awaa163

Boillée S, Yamanaka K, Lobsiger CS, Copeland NG, Jenkins NA, Kassiotis G, Kollias G, Cleveland DW. 2006. Onset and progression in inherited ALS determined by motor neurons and microglia. *Science* **312**:1389–1392.

Breiman L. 2001. Random forests. *Machine learning* **45**:5-32. doi: 10.1023/A:1010933404324

Buraschi S, Neill T, Goyal A, Iozzo RV. Decorin causes autophagy in endothelial cells via Peg3. PNAS **110**:E2582-E2591. doi: 10.1073/pnas.1305732110

Buzzao D, Castresana-Aguirre M, Guala D, Sonnhammer ELL. 2024. Benchmarking enrichment analysis methods with the disease pathway network. Brief Bioinform. **25**(2):bbae069. doi: 10.1093/bib/bbae069

Caligari M, Godi M, Guglielmetti S, Franchignoni F, Nardone A. 2013. Eye tracking communication devices in amyotrophic lateral sclerosis: impact on disability and quality of life. *Amyotroph Lateral Scler Front Degener* **14**:546–552. doi:10.3109/21678421.2013.803576

Castresana-Aguirre M, Sonnhammer ELL.2020. Pathway-specific model estimation for improved pathway annotation by network crosstalk. *Scientific Reports* **10**:13585. doi: 10.1038/s41598-020-70239-z

Comley LH, Nijssen J, Frost-Nylen J, Hedlund E. 2016. Cross-disease comparison of amyotrophic lateral sclerosis and spinal muscular atrophy reveals conservation of selective vulnerability but differential neuromuscular junction pathology. *J Comp Neurol* **524**:1424–1442. doi:10.1002/cne.23917

Elliott-Hunt CR, Marsh B, Bacon A, Pope R, Vanderplank P, Wynick D. 2004. Galanin acts as a neuroprotective factor to the hippocampus. *PNAS* **101**: 5105-5110. doi: 10.1073/pnas.0304823101.

Enjin A, Rabe N, Nakanishi ST, Vallstedt A, Gezelius H, Memic F, Lind M, Hjalt T, Tourtellotte WG, Bruder C, Eichele G, Whelan PJ, Kullander K. 2010. Identification of novel spinal cholinergic genetic subtypes disclose Chodl and Pitx2 as markers for fast motor neurons and partition cells. *J Comp Neurol* **518**: 2284-2304.

Fischer LR, Culver DG, Tennant P, David AA, Wang M, Castellano-Sanchez A, Khan J, Polak MA, Glass JD. 2004. Amyotrophic lateral sclerosis is a distal axonopathy: evidence in mice and man. *Exp Neurol* **185**: 232-240. doi: 10.1016/j.expneurol.2003.10.004

Frey D, Schneider C, Xu L, Borg J, Spooren W, Caroni, P. 2000. Early and selective loss of neuromuscular synapse subtypes with low sprouting competence in motoneuron diseases. *J Neurosci* **20**: 2534-2542. doi: 10.1523/JNEUROSCI.20-07-02534.2000.

Gautier O, Blum JA, Maksymetz J, Chen D, Schweingruber C, Mei I, Hermann A, Hackos DH, Hedlund E, Ravits J and Gitler AD. 2023. Challenges of profiling motor neuron transcriptomes from human spinal cord. *Neuron* **111**:3739-3741. doi: 10.1016/j.neuron.2023.10.035

Gizzi M, DiRocco A, Sivak M, Cohen B. 1992. Ocular motor function in motor neuron disease. *Neurology* **42**:1037–1046. doi: 10.1212/WNL.42.5.1037

Gu M, Li X, Wu R, Cheng X, Zhou S, Gu X. 2024. The transcription factor Ets1 influences axonal growth via regulation of Lcn2. *Mol. Neurobiol*. **61**: 971-981. doi: 10.1007/s12035-023-03616-0

Hänzelmann S, Castelo R, Guinney J. 2013. GSVA: gene set variation analysis for microarray and RNA-Seq data." BMC Bioinformatics, **14**, 7. doi:10.1186/1471-2105-14-7

Hor J-H, Santosa MM, Lim VJW, Ho BX, Taylor A, Khong ZJ, Ravits J, Fan Y, Liou Y-C, Soh B-S, Ng S-Y. 2021. ALS motor neurons exhibit hallmark metabolic defects that are rescued by SIRT3 activation. *Cell Death & Differentiation* **28**:1379-1397.doi:10.1038/s41418-020-00664-0

Hosack, D.A., Dennis, G., Sherman, B.T. et al. 2003. Identifying biological themes within lists of genes with EASE. *Genome Biol* **4**:P4. doi: 10.1186/gb-2003-4-6-p4

Ito Y, Yamada M, Tanaka H, Aida K, Tsuruma K, Shimazawa M, Hozumi I, Inuzuka T, Takahashi H, Hara H. 2009. Involvement of CHOP, and ER-stress apoptotic mediator in both human sporadic ALS and ALS model mice. *Neurobiol Dis* **36**:470-476. doi: 10.1016/j.nbd.2009.08.013.

Joly-Amado A, Hunter J, Quadri Z, Zamudio F, Rocha-Rangel PV, Chan D, Kesarwani A, Nash K, Lee DC, Morgan D, Gordon MN, Selenica MLB. 2020. Ccl2 overexpression in the brain promotes glial activation and accelerates tau pathology in a mouse model of tauopathy. *Front Immunol* **11**:997. doi: 10.3389/fimmu.2020.00997

Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A. 2021. Fast gene set enrichment analysis. *bioRxiv* doi: 10.1101/060012

Kubota M, Sakakihara Y, Uchiyama Y, Nara A, Nagata T, Nitta H, Ishimoto K, Oka A, Horio K, Yanagisawa M. 2000. New ocular movement detector system as a communication tool in ventilator-assisted Werdnig–Hoffmann disease. *Dev Med Child Neurol* **42**:61–64. doi:

10.1017/s0012162200000116

Leboeuf M, Vargas-Abonce SE, Pezé-Hedsieck E, Dupont E, Jimenez-Alonso L, Moya KL, Prochiantz A. 2023. ENGRAILED-1 transcription factor has a paracrine neurotrophic activity on adult spinal α-motorneurons. *EMBO Reports* **24**:e56525 doi: 10.15252/embr.202256525

Liu S, Costa M. 2022. The role of NUPR1 in response to stress and cancer development. *Toxicol Appl Pharmacol* **454**: 116244. doi: 10.1016/j.taap.2022.116244

Lobsiger CS, Boillee S, Cleveland DW. 2007. Toxicity from different SOD1 mutants dysregulates the complement system and the neuronal regenerative response in ALS motor neurons. *PNAS* **104**: 7319-7326. doi: 10.1073/pnas.0702230104

Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**:550. doi: 10.1186/s13059-014-0550-8

Mehta AR, Gregory JM, Dando O, Carter RN, Burr K, Nanda J, Story D, McDade K, Smith C, Morton NM, Mahad DJ, Hardingham GE, Chandran S, Selvaraj BT. 2021. Mitochondrial bioenergetic deficits in C9orf72 amyotrophic lateral sclerosis motor neurons cause dysfunctional axonal homeostasis. *Acta Neuropathol.* **141**:257-279. doi:10.1007/s00401-020-02252-5

Müller D, Cherukuri P, Henningfeld K, Poh CH, Wittler L, Grote P, Schlüter O, Schmidt J, Laborda J, Bauer SR, Brownstone RM, Marquardt T. 2014. Dlk1 promotes a fast motor neuron biophysical signature required for peak force execution. *Science* **343**:1264-1266. doi:

10.1126/science.1246448

Namboori SC, Thomas P, Ames R, Hawkins S, Garrett LO, Willis CRG, Rosa A, Stanton LW, Bhinge A. 2021. Single-cell transcriptomics identifies master regulators of neurodegeneration in SOD1 ALS iPSC-derived motor neurons. *Stem Cell Reports* **16**: 3020-3035. doi: 10.1016/j.stemcr.2021.10.010

Nichterwitz S, Chen G, Aguila Benitez J, Yilmaz M, Storvall H, Cao M, Sandberg R, Deng Q, Hedlund E. 2016. Laser capture microscopy coupled with Smart-seq2 (LCM-seq) for robust and efficient transcriptomic profiling of mouse and human cells. *Nat Commun* **7**: 12139. doi:10.1038/ncomms12139

Nichterwitz S, Benitez JA, Hoogstraaten R, Deng Q, Hedlund E. 2018. LCMseq: a method for spatial transcriptomic profiling using laser capture microdissection coupled with polyA-based RNA sequencing. *Methods Mol Biol* **1649**: 95–110. doi:10.1007/978-1-4939-7213-5_6

Nichterwitz S, Nijssen J, Storvall H, Schweingruber C, Comley LH, Allodi I, van der Lee M, Deng Q, Sandberg R and Hedlund E. 2020. LCM-seq reveals unique transcriptional adaptation mechanisms and protective pathways of resistant neurons in spinal muscular atrophy. *Genome Research* **30**:1083-1096. doi: 10.1101/gr.265017.120

Nijssen J, Comley LK, Hedlund E. 2017. Motor neuron vulnerability and resistance in amyotrophic

lateral sclerosis. *Acta Neuropathol* **133**: 863-885. doi: 10.1007/s00401-017-1708-8

Pasinetti GM, Ungar LH, Lange DJ, Yemul S, Deng H, Yuan X, Brown RH, Cudkowicz ME, Newhall K, Peskind E, Marcus S, Ho L. 2006. Identification of potential CSF biomarkers in ALS. *Neurology* **66**: 1218-1222. doi: 10.1212/01.wnl.0000203129.82104.07

Pedersen WA, Wan R, Zhang P, Mattson MP. 2002. Urocortin, but not Urocortin II, protects cultured hippocampal neurons from oxidative and excitotoxic cell death via corticotropin-releasing hormone receptor type I. *J Neurosci* **22**:404-412. doi: 10.1523/JNEUROSCI.22-02-00404.2002

Perrin FE, Boisset G, Docquier M, Schaad O, Descombes P, Kato AC. 2005. No widespread induction of cell death genes occurs in pure motoneurons in an amyotrophic lateral sclerosis mouse model. *Hum Mol Genet* **14**: 3309-3320. doi: 10.1093/hmg/ddi357.

Pun S, Santos AF, Saxena S, Xu L, Caroni P. 2006. Selective vulnerability and pruning of phasic motoneuron axons in motoneuron disease alleviated by CNTF. *Nat Neurosci* **9**: 408-419. doi:10.1038/nn1653

Ries A, Goldberg JL, Grimpe B. 2007. A novel biological function for CD44 in axon growth of retinal ganglion cells identified by a bioinformatics approach. *J Neurochem* **103**:1491-1505. doi: 10.1111/j.1471-4159.2007.04858.x

Robinson MD, McCarthy DJ and Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**: 139-140. doi: 10.1093/bioinformatics/btp616

Rosen DR, Siddique T, Patterson D, Figlewicz DA, Sapp P, Hentati A, Donaldson D, Goto J, O'Regan JP, Deng H-X, Rahmani Z, Krizus A, McKenna-Yasek D, Cayabyab A, Gaston SM, Berger R, Tanzi RE, Halperin JJ, Herzfeldt B, Van den Bergh R, Hung W-Y, Bird T, Deng G, Mulder DW, Smyth C, Laing NG, Soriano E, Pericak-Vance MA, Haines J, Rouleau GA, Gusella JS, Horvitz HR, Brown Jr. RH. 1993. Mutations in Cu/Zn superoxide dismutase gene are associated with familial amyotrophic lateral sclerosis. *Nature* **362**: 59-62. doi: 10.1038/362059a0

Saxena S, Cabuy E, Caroni P. 2009. A role for motoneuron subtype-selective ER stress in disease manifestations of FALS mice. *Nat Neurosci* **12**: 627-636. doi: 10.1038/nn.2297.

Schaefer AM, Sanes JR, Lichtman JW. 2005. A compensatory subpopulation of motor neurons in a mouse model of amyotrophic lateral sclerosis. *J Comp Neurol.* **490**: 209-219. doi: 10.1002/cne.20620

Schweingruber C, Nijssen J, Mechtersheimer J, Reber S, Keuper M, Mei I, Agulia J, O'Brien N, Jastroch M, Ruepp MD and Hedlund E. 2023. Single cell RNA sequencing in isogenic FUS and TARBDP mutant ALS lines reveals early mitochondrial dysfunction as a common pathway specific

to motor neurons. *bioRxiv*. doi: 10.1101/2023.03.16.531876v2

Seijffers R, Zhang J, Matthews JC, Chen A, Tamrazian E, Babaniyi O, Selig M, Hynynen M, Woolf CJ, Brown Jr RH. 2014. ATF3 expression Improves motor function in the ALS Mouse model by promoting motor neuron survival and retaining muscle innervation. *PNAS* **111**:1622-1627. doi: 10.1073/pnas.1314826111

Shadrach JL, Stansberry WM, Milen AM, Fogarty EA, Antonellis A. 2021. Translatomic analysis of regenerating and degenerating spinal motor neurons in injury and ALS. *iScience* **24**: 102700. doi: 10.1016/j.isci.2021.102700

Shih CH, Lacagnina M, Leuer-Bisciotti K, Pröschel C. 2014. Astroglial-derived periostin promotes axonal regeneration after spinal cord injury. *J Neurosci* **34**: 2438-2443. doi: 10.1523/JNEUROSCI.2947-13.2014

Shimamura M, Taniyama Y, Katsuragi N, Koibuchi N, Kyutoku M, Sato N, Allahtavakoli M, Wakayama K, Nakagami H, Morishita R. 2012. Role of central nervous system periostin in cerebral ischemia. *Stroke* **43**: 1108-1114. doi: 10.11617STROKEAHA.111.636662.

Shimizu T, Hayashi M, Kawata A, Mizutani T, Watabe K, Matsubara S (2011) A morphometric study of the vagus nerve in amyotrophic lateral sclerosis with circulatory collapse. *Amyotrophic Lateral Sclerosis* **12**:356-362. doi: 10.3109/17482968.2011.566342

Spiller KJ, Cheung CJ, Restrepo CR, Kwong LK, Stieber AM, Trojanowski JQ, Lee VM (2016) Selective motor neuron resistance and recovery in a new inducible mouse model of TDP-43 proteinopathy. *J Neurosci* **36**:7707–7717. doi:10.1523/jneurosci.1457-16.2016

Starkey ML, Davies M, Yip PK, Carter LM, Wong DJN, McMahon SB, Bradbury EJ. 2012. Expression of the regeneration-associated protein SPRR1A in primary sensory neurons and spinal cord of the adult mouse following peripheral and central injury. *J Comp Neurol* **513**:51-68. doi: 10.1002/cne.21944

Stirling DR, Swain-Bowden MJ, Lucas AM, Carpenter AE, Cimini BA, Goodman A. 2021. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinformatics* **22**: 433. doi: 10.1186/s12859-021-04344-9

Sun S, Sun Y, Ling S-C, Ferraiuolo L, McAlonis-Downes M, Zou Y, Drenner K, Wang Y, Ditsworth D, Tokunaga S, Kopelevich A, Kaspar BK, Lagier-Tourenne C, Cleveland DW. 2015. Translational profiling identifies a cascade of damage initiated in motor neurons and spreading to glia in mutant SOD1-mediated ALS. *Proc Natl Acad Sci USA* **112**: E6993-E7002. doi: 10.1073/pnas.1520639112

Tanaka H, Shimazawa M, Kimura M, Takata M, Tsuruma K, Yamada M, Takahashi H, Hozumi I, Niwa J-I, Iguchi Y, Nikawa T, Sobue G, Inuzuka T, Hara H. 2012. The potential of GPNMB as novel

neuroprotective factor in amyotrophic lateral sclerosis. *Scientific Reports* **2**: 573 doi: 10.1038/srep00573

Tezze C, Romanello V, Sandri M. 2019. FGF21 as modulator of metabolism in health and disease. *Front Physiol*. **10**: 419 doi: 10.3389/fphys.2019.00419

Thomas EV, Nagy M, Zhao H, Fenton WA, Horwich AL. 2018. Motor nuclei innervating eye muscles spared in a mouse model of SOD1-linked ALS. *bioRxiv*. doi: 10.1101/304857

Tjust AE, Brannstrom T, Pedrosa-Domellof F. 2012. Unaffected motor endplate occupancy in eye muscles of ALS G93A mouse model. *Front Biosci (Schol Ed)* **4**:1547–1555. doi: 10.2741/s351

Valdez G, Tapia JC, Lichtman JW, Fox MA, Sanes JR. 2012. Shared resistance to aging and ALS in neuromuscular junctions of specific muscles. *PLoS One* **7**:e34640. doi: 10.1371/journal.pone.0034640

Wozniak EAL, Chen Z, Paul S, Yang P, Figueroa KP, Friedrich J, Tschumperlin T, Berken M, Ingram M, Henzler C, Pulst SM, Orr HT. 2021. Cholecystokinin 1 receptor activation restores normal mTORC1 signaling and is protective to Purkinje cells of SCA mice. *Cell Rep* **37**:109831. doi: 10.1016/j.celrep.2021.109831

Van Den Bosch L, Schwaller B, Vleminckx V, Meijers B, Stork S, Ruehlicke T, Van Houtte E, Klaassen

H, Celio MR, Missiaen L, Robberecht W, Berchtold MW. 2002. Protective effect of parvalbumin on excitotoxic motor neuron death. *Exp Neurol* **174**:150-161. doi: 10.1006/exnr.2001.7858.

Yamanaka K, Chun SJ, Boillee S, Fujimori-Tonou N, Yamashita H, Gutmann DH, Takahashi R, Misawa H, Cleveland DW. 2008. Astrocytes as determinants of disease progression in inherited amyotrophic lateral sclerosis. *Nat Neurosci* **11**:251–253. doi: 10.1038/nn2047

Zhang Y, Cook A, Kim Jinho, Baranov SV, Jiang J, SMith K, Cormier K, Bennett E, Browser RP, Day AL, Carlisle D, Ferrante RJ, Wang X, Friedlander RM. 2013. melatonin inhibits the caspase-1/cytochromec/caspase-3 cell death pathway, inhibits MT1 receptor loss and delays disease progression in a mouse model of amyotrophic lateral sclerosis. *Neurobiol Dis* **55**:26-35. doi: 10.1016/j.nbd.2013.03.008

Zhang Y, Chen K, SLoan SA, Bennett ML, Scholze AR, O'Keeffe S, Phatnani HP, Guarnieri P, Caneda C, Ruderisch N, et al. 2014. An RNA-sequencing transcriptome and splicing database of glia, neurons, and vascular cells of the cerebral cortex. *J Neurosci* **34**: 11929-11947. doi:10.1523/JNEUROSCI 1860-14.2014

Zhao Z, Lange DJ, Ho L, Bonini S, Shao B, Salton SR, Thomas S, Pasinetti GM. 2008. Vgf is a novel biomarker associated with muscle weakness in amyotrophic lateral sclerosis (ALS), with a potential role in disease pathogenesis. *Int J Med Sci* **5**: 92-99. doi: 10.7150/ijms.5.92

Zhou Y, Tang H, Liu J, Dong J, Ziong H. 2011. Chemokine CCL2 modulation of neuronal excitability and synaptic transmission in rat hippocampal slices. *J Neurochem* **116**:406-414. doi: 10.1111/j.1471-4159-2010.07121.x

Figure legends

Fig. 1 Motor neuron subpopulations show unique transcriptional profiles but similar levels of ALS disease gene expression. A. Scheme of LCM-seq: we used G93A hSOD1 mice as a model for ALS and wildtype littermates as controls at the age of P56 and P112. **B**. Pairwise Pearson's correlation on vst (variance standard transformation) normalized counts of all samples per cell type, genotype and age. Single linkage method was used for the hierarchical clustering of columns. **C**. Principal component analysis on the whole gene expression data set. **D**. Heatmap of $\log_2$ normalized rpkm expression of ALS genes. (P) Postnatal day; (LCM-seq) laser capture microdissection coupled with RNA sequencing; (SC) spinal cord; (CN) cranial nerve; (CN12) hypoglossal nucleus; (CN10) dorsal motor nucleus of the vagus nerve; (CN3/4) oculomotor and trochlear nuclei.

Fig.2. Resistant motor neurons show little gene dysregulation in response to mutant SOD1

(**A**). Bar plots of up- and downregulated DEGs for all motor neuron populations. (**B,C**) Upset plot of DEGs at (**B**) P56 and (**C**) P112 for all motor neuron populations (Likelihood ratio test, Benjamini Hochberg adjusted p value < 0.05). (**D**) Heatmap showing logFC expression of the DEGs across disease stages for CN3/4. (Likelihood ratio test, ** Benjamini Hochberg FDR < 0.01; * Benjamini

Hochberg FDR < 0.05). (**E**) Expression of Dlk1 across cell types and genotypes. (**F)** Heatmaps showing logFC expression of the DEGs across disease stages for CN10. (Likelihood ratio test, ** Benjamini Hochberg FDR < 0.01; * Benjamini Hochberg FDR < 0.05).

Fig. 3. Baseline gene expression in CN3/4 motor neurons may hold the key to their resilience

(**A**) Upset plot of detected genes in CN3/4 and SC in WT mice at P56 and P112. (**B**) Volcano plot showing genes significantly enriched in SC (light blue) or CN3/4 (light purple) at P56 (Likelihood ratio test, Benjamini Hochberg adjusted p value < 0.05). (**C**) Volcano plot showing genes significantly enriched in SC (blue) or CN3/4 (purple) at P112 (Likelihood ratio test, Benjamin Hochberg adjusted p value < 0.05). (**D**) GO terms enriched in WT CN3/4 or WT SC motor neurons. (**E-G**) Selected pathways enriched in CN3/4 in wild-type mice.

Fig. 4. Vulnerable motor neurons show unique regulation of injury response genes.

(**A**) Upset plot of DEGs in vulnerable population at P56 and P112 (Likelihood ratio test, Benjamini Hochberg adjusted p value < 0.05). (**B**) Heatmap showing logFC expression of common DEGs between ages in CN12. (**C**) Heatmap showing logFC expression of common DEGs in CN12 and SC at P112. (**D**) Heatmap showing logFC expression of common DEGs between ages in spinal motor neurons. (**E**) Heatmap showing logFC expression of DEGs related to inflammation. (**B-E**) Likelihood ratio test, ** Benjamini Hochberg FDR < 0.01; * Benjamini Hochberg FDR < 0.05. (**F-L**) Representative RNAscope images of Atf3, Fgf21, Timp1, Sprr1a, and Vgf (scale bars, 50 µm) (Scale bars: 30 µm, permutation test **** P<0.00001).

Fig. 5. Pathway analysis shows a common detrimental response across vulnerable populations. (**A,B**) GO term dot plot showing the GO terms that were enriched in at least two of the pathway analysis methods (Fisher test, fGSEA, Anubix) for CN12 and SC at P112. Enrichment score correspond to the amount of functional genes that the method shows being related for the enrichment term. (**C**) Sankey plot showing how stressed genes belong to each relevant pathway in CN12 and SC at P112. (**D**) Network plot of Funcoup evidence types with vulnerable genes shared between CN12 and SC at P112. (**E**) PCA plot illustrates the clustering pattern of samples based on DEGs belonging to common detrimental pathways ('Positive regulation of cell death', 'Inflammatory response', 'Response to wounding', 'Negative regulation of cell death', 'ERK1 and ERK2 cascade') for CN12 and SC cell types at P112. Top 5 loading genes for PC1 are shown in the plot.

Fig. 6. Dataset comparison and feature selection show a common importance for *Vgf* in ALS. (**A**) Overview of the DEGs comparison between Namboori et al. 2021 and SC (P112) dataset (**B**) Schematic of the Machine learning approach on gene expression data. (**C**) Features importance obtained from random forest approach. (**D**) Overview of the datasets we used for a cross comparison of DEGs: Shadrach et al. 2021, Lobsiger et al. 2007, Sun et al. 2015, Namboori et al. 2021. (**E**) Upset plot of DEGs for all datasets (Likelihood ratio test, Benjamini Hochberg adjusted p value < 0.05). (**F**) Heatmap showing the common DEGs between Shadrach et al. 2021, Lobsiger et al. 2007, Namboori et al. 2021 and SC at symptomatic stage. (**G**) Heatmap showing common DEGs between all datasets. (**H**) Representative immunofluorescence image of SC at P112 and

showing the expression of VGF protein in VAChT positive cells, Scale bar: 30 µm.

.

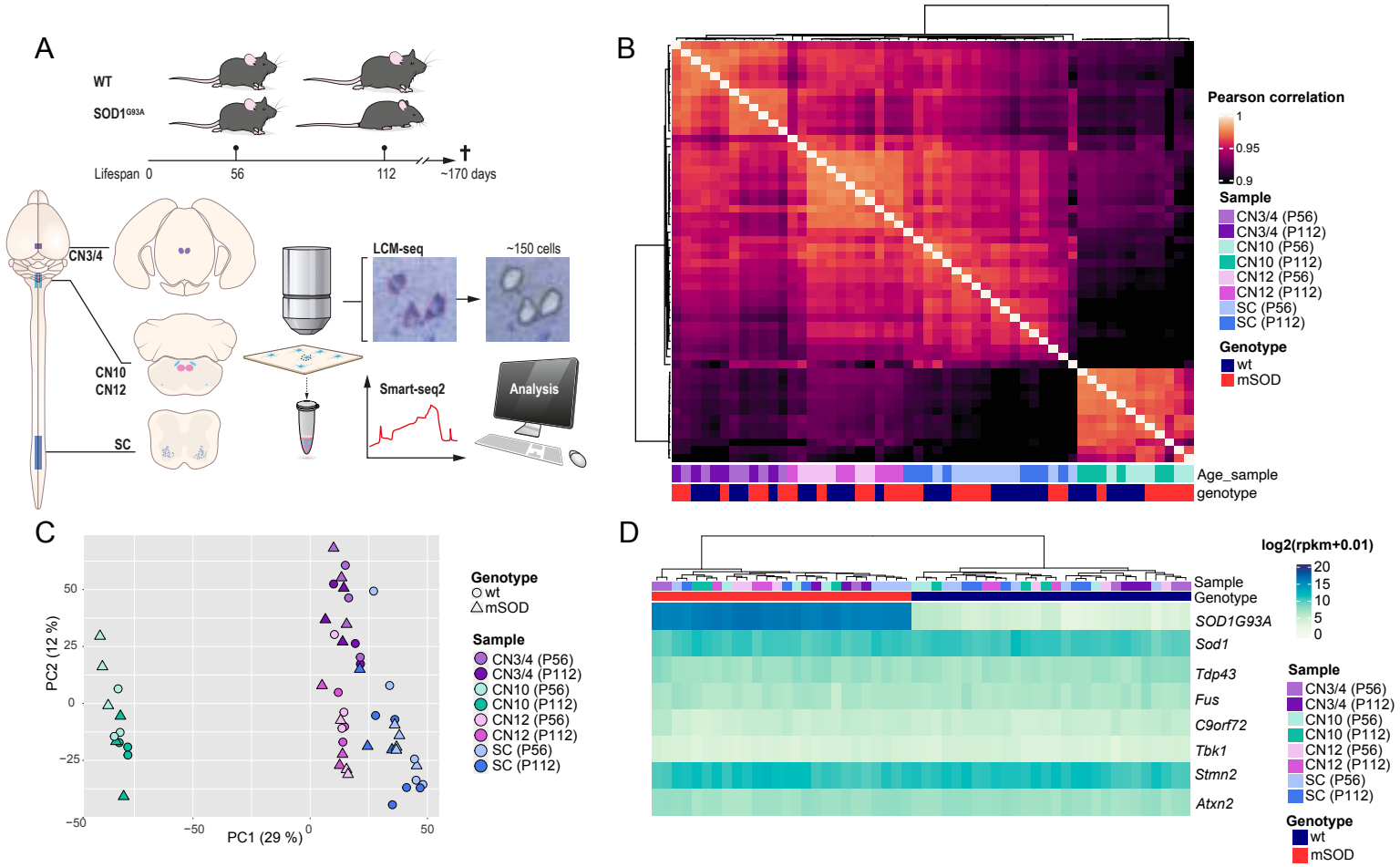
A
WT
SOD1G93A
Lifespan 0
56
112
~170 days
CN3/4
CN10
CN12
SC
LCM-seq
~150 cells
Smart-seq2
Analysis
B
Pearson correlation
1
0.95
0.9
Sample
CN3/4 (P56)
CN3/4 (P112)
CN10 (P56)
CN10 (P112)
CN12 (P56)
CN12 (P112)
SC (P56)
SC (P112)
Genotype
wt
mSOD
Age_sample
genotype
C
PC2 (12 %)
PC1 (29 %)
50
25
0
−25
−50
Genotype
wt
mSOD
Sample
CN3/4 (P56)
CN3/4 (P112)
CN10 (P56)
CN10 (P112)
CN12 (P56)
CN12 (P112)
SC (P56)
SC (P112)
D
log2(rpkm+0.01)
20
15
10
5
0
Sample
Genotype
SOD1G93A
Sod1
Tdp43
Fus
C9orf72
Tbk1
Stmn2
Atxn2
Sample
CN3/4 (P56)
CN3/4 (P112)
CN10 (P56)
CN10 (P112)
CN12 (P56)
CN12 (P112)
SC (P56)
SC (P112)
Genotype
wt
mSOD

Figure 1

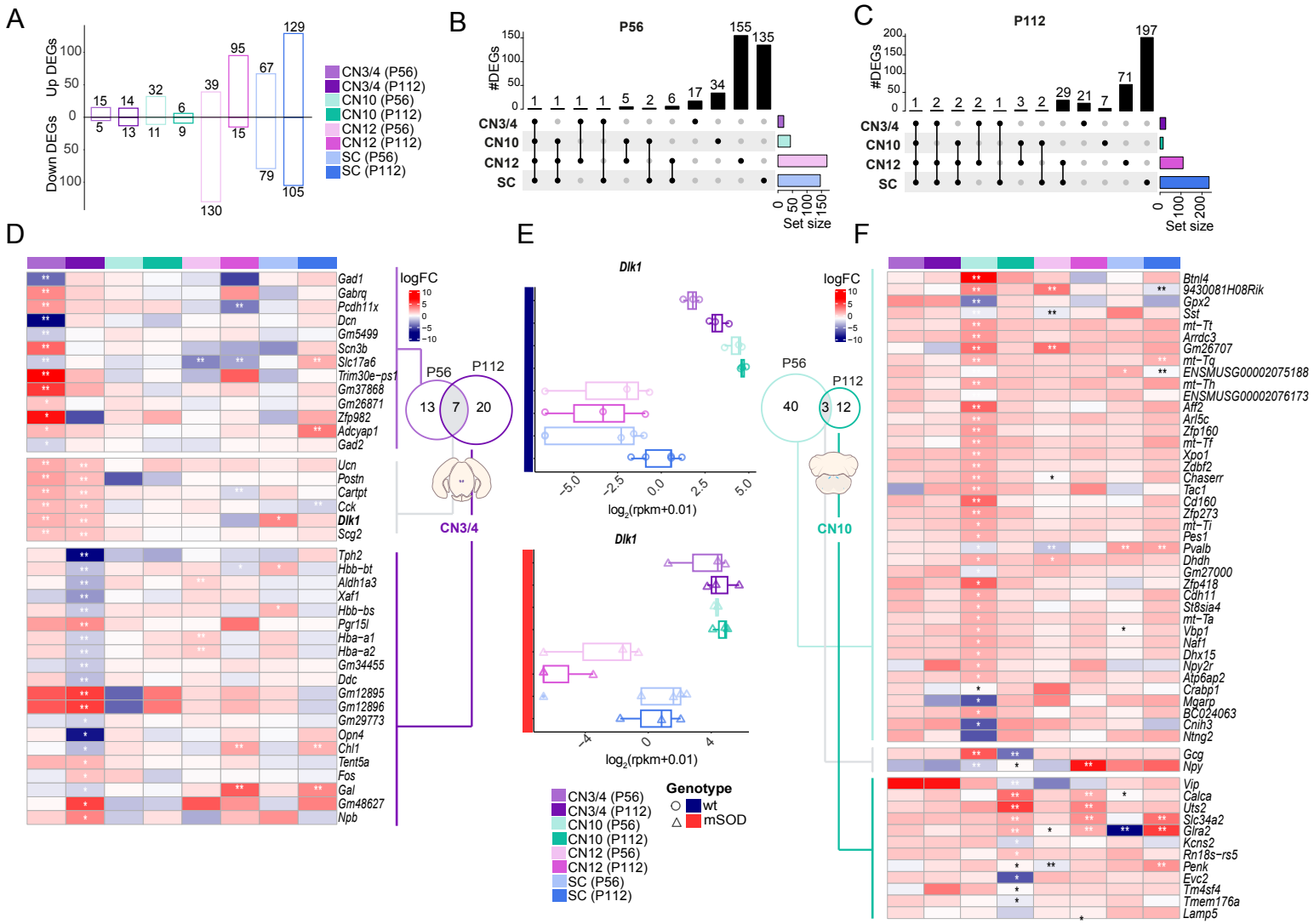
A
Up DEGs
Down DEGs
CN3/4 (P56)
CN3/4 (P112)
CN10 (P56)
CN10 (P112)
CN12 (P56)
CN12 (P112)
SC (P56)
SC (P112)
B
P56
#DEGs
CN3/4
CN10
CN12
SC
Set size
C
P112
D
logFC
P56
P112
CN3/4
E
Dlk1
log2(rpkm+0.01)
Genotype
wt
mSOD
F
CN10

Figure 2

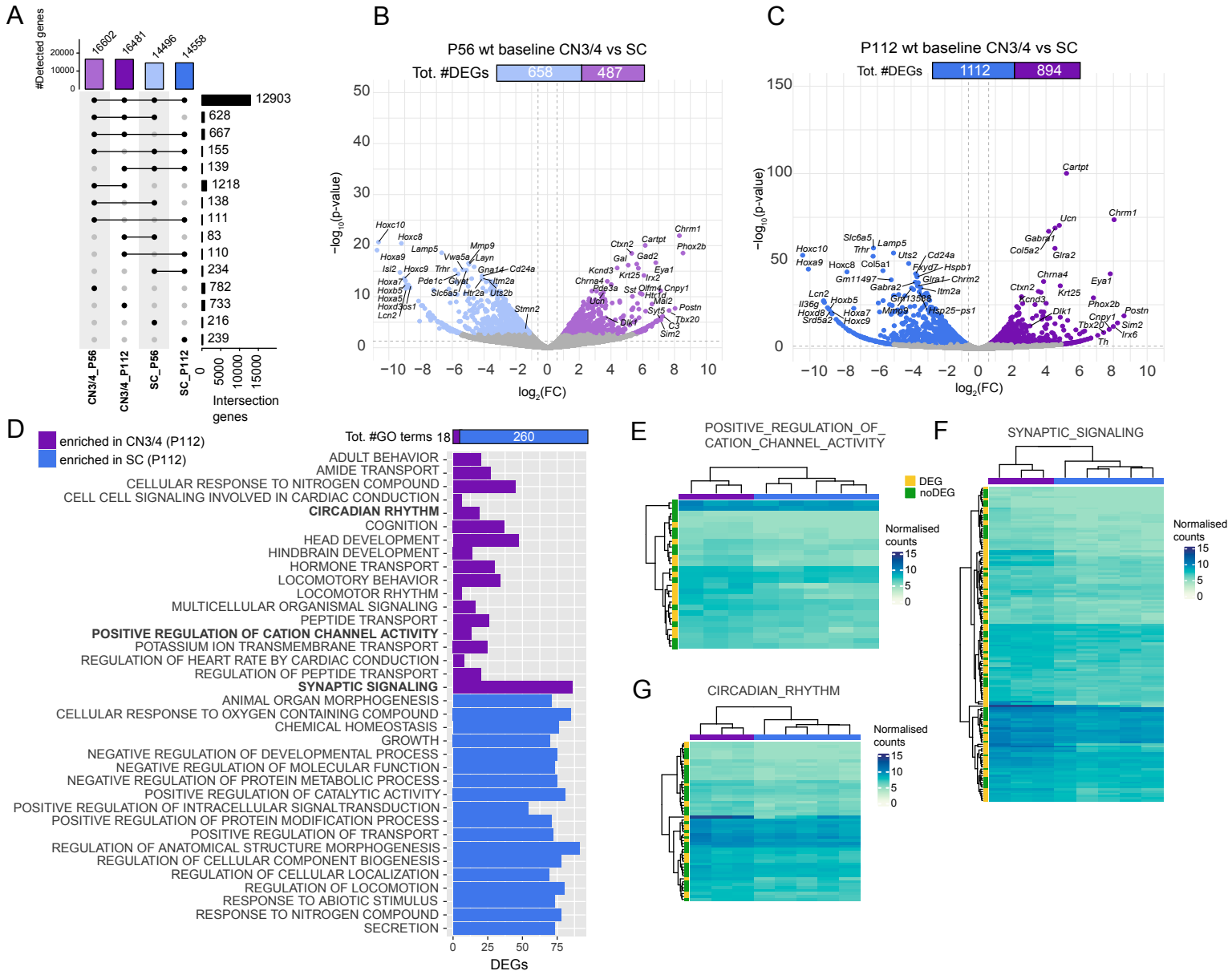
A
#Detected genes
16602
16481
14496
14558
12903
628
667
155
139
1218
138
111
83
110
234
782
733
216
239
CN3/4_P56
CN3/4_P112
SC_P56
SC_P112
Intersection genes
B
P56 wt baseline CN3/4 vs SC
Tot. #DEGs
658
487
−log10(p-value)
log2(FC)
C
P112 wt baseline CN3/4 vs SC
Tot. #DEGs
1112
894
−log10(p-value)
log2(FC)
D
enriched in CN3/4 (P112)
enriched in SC (P112)
Tot. #GO terms 18
260
ADULT BEHAVIOR
AMIDE TRANSPORT
CELLULAR RESPONSE TO NITROGEN COMPOUND
CELL CELL SIGNALING INVOLVED IN CARDIAC CONDUCTION
CIRCADIAN RHYTHM
COGNITION
HEAD DEVELOPMENT
HINDBRAIN DEVELOPMENT
HORMONE TRANSPORT
LOCOMOTORY BEHAVIOR
LOCOMOTOR RHYTHM
MULTICELLULAR ORGANISMAL SIGNALING
PEPTIDE TRANSPORT
POSITIVE REGULATION OF CATION CHANNEL ACTIVITY
POTASSIUM ION TRANSMEMBRANE TRANSPORT
REGULATION OF HEART RATE BY CARDIAC CONDUCTION
REGULATION OF PEPTIDE TRANSPORT
SYNAPTIC SIGNALING
ANIMAL ORGAN MORPHOGENESIS
CELLULAR RESPONSE TO OXYGEN CONTAINING COMPOUND
CHEMICAL HOMEOSTASIS
GROWTH
NEGATIVE REGULATION OF DEVELOPMENTAL PROCESS
NEGATIVE REGULATION OF MOLECULAR FUNCTION
NEGATIVE REGULATION OF PROTEIN METABOLIC PROCESS
POSITIVE REGULATION OF CATALYTIC ACTIVITY
POSITIVE REGULATION OF INTRACELLULAR SIGNALTRANSDUCTION
POSITIVE REGULATION OF PROTEIN MODIFICATION PROCESS
POSITIVE REGULATION OF TRANSPORT
REGULATION OF ANATOMICAL STRUCTURE MORPHOGENESIS
REGULATION OF CELLULAR COMPONENT BIOGENESIS
REGULATION OF CELLULAR LOCALIZATION
REGULATION OF LOCOMOTION
RESPONSE TO ABIOTIC STIMULUS
RESPONSE TO NITROGEN COMPOUND
SECRETION
DEGs
E
POSITIVE_REGULATION_OF_CATION_CHANNEL_ACTIVITY
DEG
noDEG
Normalised counts
F
SYNAPTIC_SIGNALING
Normalised counts
G
CIRCADIAN_RHYTHM
Normalised counts

Figure 3

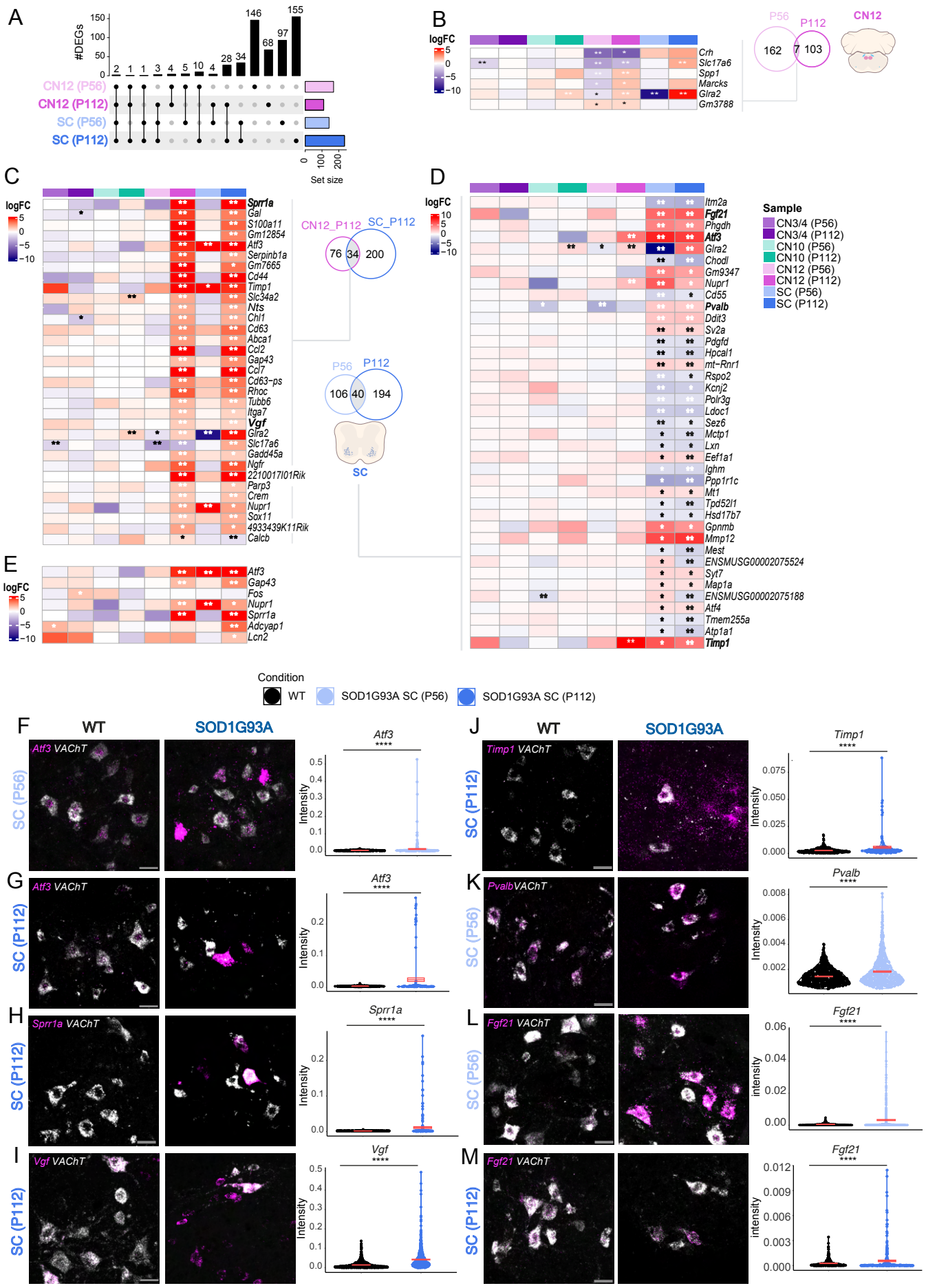
A
#DEGs
Set size
CN12 (P56)
CN12 (P112)
SC (P56)
SC (P112)
B
logFC
Crh
Slc17a6
Spp1
Marcks
Glra2
Gm3788
P56
P112
CN12
162
7
103
C
logFC
Sprr1a
Gal
S100a11
Gm12854
Atf3
Serpinb1a
Gm7665
Cd44
Timp1
Slc34a2
Nts
Chl1
Cd63
Abca1
Ccl2
Gap43
Ccl7
Cd63-ps
Rhoc
Tubb6
Itga7
Vgf
Glra2
Slc17a6
Gadd45a
Ngfr
2210017I01Rik
Parp3
Crem
Nupr1
Sox11
4933439K11Rik
Calcb
CN12_P112
SC_P112
76
34
200
P56
P112
106
40
194
SC
D
logFC
Itm2a
Fgf21
Phgdh
Atf3
Glra2
Chodl
Gm9347
Nupr1
Cd55
Pvalb
Ddit3
Sv2a
Pdgfd
Hpcal1
mt-Rnr1
Rspo2
Kcnj2
Polr3g
Ldoc1
Sez6
Mctp1
Lxn
Eef1a1
Ighm
Ppp1r1c
Mt1
Tpd52l1
Hsd17b7
Gpnmb
Mmp12
Mest
ENSMUSG00002075524
Syt7
Map1a
ENSMUSG00002075188
Atf4
Tmem255a
Atp1a1
Timp1
Sample
CN3/4 (P56)
CN3/4 (P112)
CN10 (P56)
CN10 (P112)
CN12 (P56)
CN12 (P112)
SC (P56)
SC (P112)
E
logFC
Atf3
Gap43
Fos
Nupr1
Sprr1a
Adcyap1
Lcn2
Condition
WT
SOD1G93A SC (P56)
SOD1G93A SC (P112)
F
WT
SOD1G93A
SC (P56)
Atf3 VAChT
Atf3
Intensity
G
SC (P112)
Atf3 VAChT
Atf3
H
SC (P112)
Sprr1a VAChT
Sprr1a
I
SC (P112)
Vgf VAChT
Vgf
J
WT
SOD1G93A
SC (P112)
Timp1 VAChT
Timp1
K
SC (P56)
PvalbVAChT
Pvalb
L
SC (P56)
Fgf21 VAChT
Fgf21
M
SC (P112)
Fgf21 VAChT
Fgf21

Figure 4

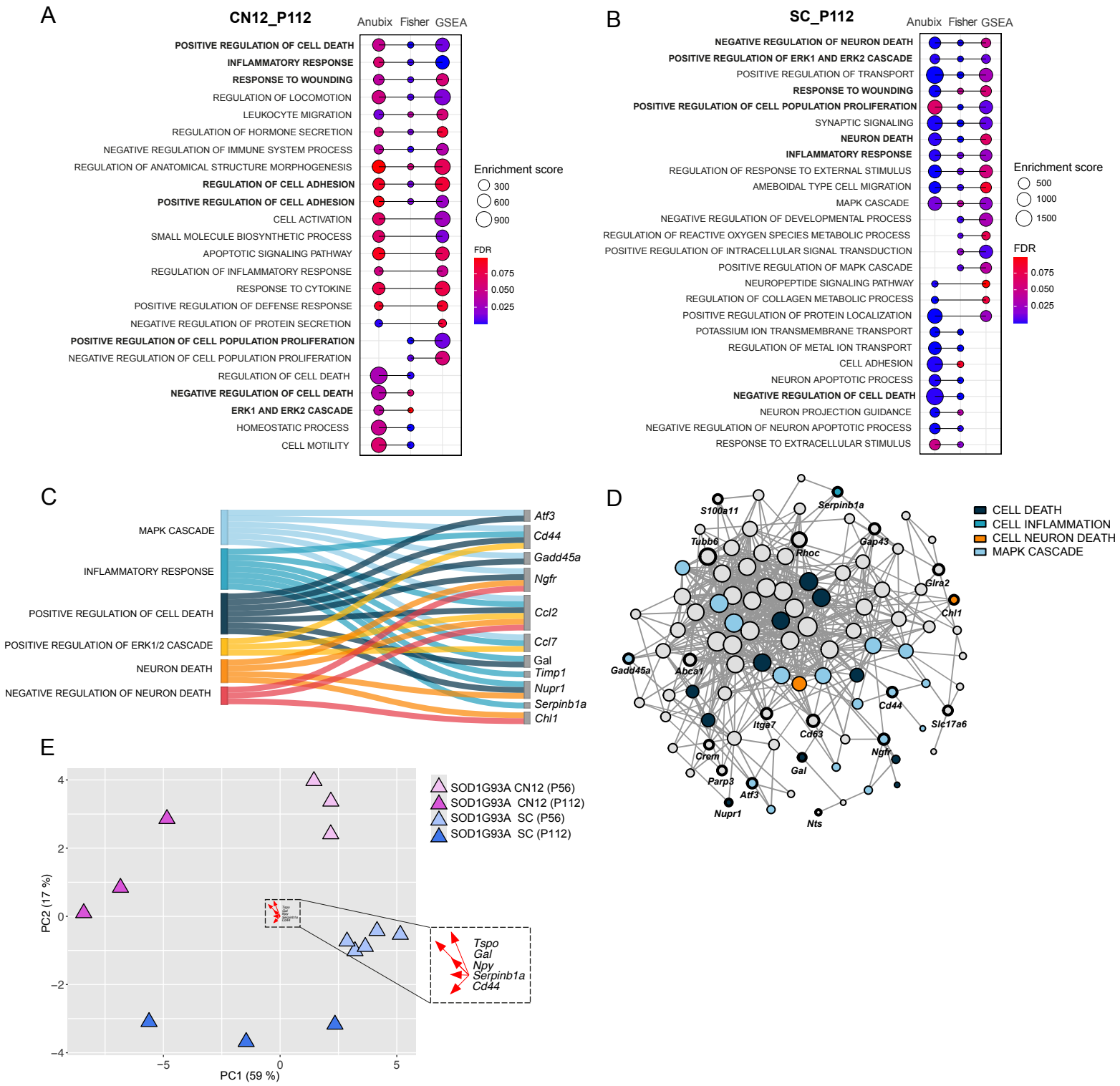
A
CN12_P112
Anubix Fisher GSEA
POSITIVE REGULATION OF CELL DEATH
INFLAMMATORY RESPONSE
RESPONSE TO WOUNDING
REGULATION OF LOCOMOTION
LEUKOCYTE MIGRATION
REGULATION OF HORMONE SECRETION
NEGATIVE REGULATION OF IMMUNE SYSTEM PROCESS
REGULATION OF ANATOMICAL STRUCTURE MORPHOGENESIS
REGULATION OF CELL ADHESION
POSITIVE REGULATION OF CELL ADHESION
CELL ACTIVATION
SMALL MOLECULE BIOSYNTHETIC PROCESS
APOPTOTIC SIGNALING PATHWAY
REGULATION OF INFLAMMATORY RESPONSE
RESPONSE TO CYTOKINE
POSITIVE REGULATION OF DEFENSE RESPONSE
NEGATIVE REGULATION OF PROTEIN SECRETION
POSITIVE REGULATION OF CELL POPULATION PROLIFERATION
NEGATIVE REGULATION OF CELL POPULATION PROLIFERATION
REGULATION OF CELL DEATH
NEGATIVE REGULATION OF CELL DEATH
ERK1 AND ERK2 CASCADE
HOMEOSTATIC PROCESS
CELL MOTILITY
Enrichment score
300
600
900
FDR
0.075
0.050
0.025
B
SC_P112
Anubix Fisher GSEA
NEGATIVE REGULATION OF NEURON DEATH
POSITIVE REGULATION OF ERK1 AND ERK2 CASCADE
POSITIVE REGULATION OF TRANSPORT
RESPONSE TO WOUNDING
POSITIVE REGULATION OF CELL POPULATION PROLIFERATION
SYNAPTIC SIGNALING
NEURON DEATH
INFLAMMATORY RESPONSE
REGULATION OF RESPONSE TO EXTERNAL STIMULUS
AMEBOIDAL TYPE CELL MIGRATION
MAPK CASCADE
NEGATIVE REGULATION OF DEVELOPMENTAL PROCESS
REGULATION OF REACTIVE OXYGEN SPECIES METABOLIC PROCESS
POSITIVE REGULATION OF INTRACELLULAR SIGNAL TRANSDUCTION
POSITIVE REGULATION OF MAPK CASCADE
NEUROPEPTIDE SIGNALING PATHWAY
REGULATION OF COLLAGEN METABOLIC PROCESS
POSITIVE REGULATION OF PROTEIN LOCALIZATION
POTASSIUM ION TRANSMEMBRANE TRANSPORT
REGULATION OF METAL ION TRANSPORT
CELL ADHESION
NEURON APOPTOTIC PROCESS
NEGATIVE REGULATION OF CELL DEATH
NEURON PROJECTION GUIDANCE
NEGATIVE REGULATION OF NEURON APOPTOTIC PROCESS
RESPONSE TO EXTRACELLULAR STIMULUS
Enrichment score
500
1000
1500
FDR
0.075
0.050
0.025
C
MAPK CASCADE
INFLAMMATORY RESPONSE
POSITIVE REGULATION OF CELL DEATH
POSITIVE REGULATION OF ERK1/2 CASCADE
NEURON DEATH
NEGATIVE REGULATION OF NEURON DEATH
Atf3
Cd44
Gadd45a
Ngfr
Ccl2
Ccl7
Gal
Timp1
Nupr1
Serpinb1a
Chl1
D
CELL DEATH
CELL INFLAMMATION
CELL NEURON DEATH
MAPK CASCADE
S100a11
Serpinb1a
Tubb6
Rhoc
Gap43
Glra2
Chl1
Gadd45a
Abca1
Cd44
Slc17a6
Itga7
Cd63
Ngfr
Crem
Gal
Parp3
Atf3
Nupr1
Nts
E
SOD1G93A CN12 (P56)
SOD1G93A CN12 (P112)
SOD1G93A SC (P56)
SOD1G93A SC (P112)
PC2 (17 %)
PC1 (59 %)
4
2
0
−2
−4
−5
0
5
Tspo
Gal
Npy
Serpinb1a
Cd44

Figure 5

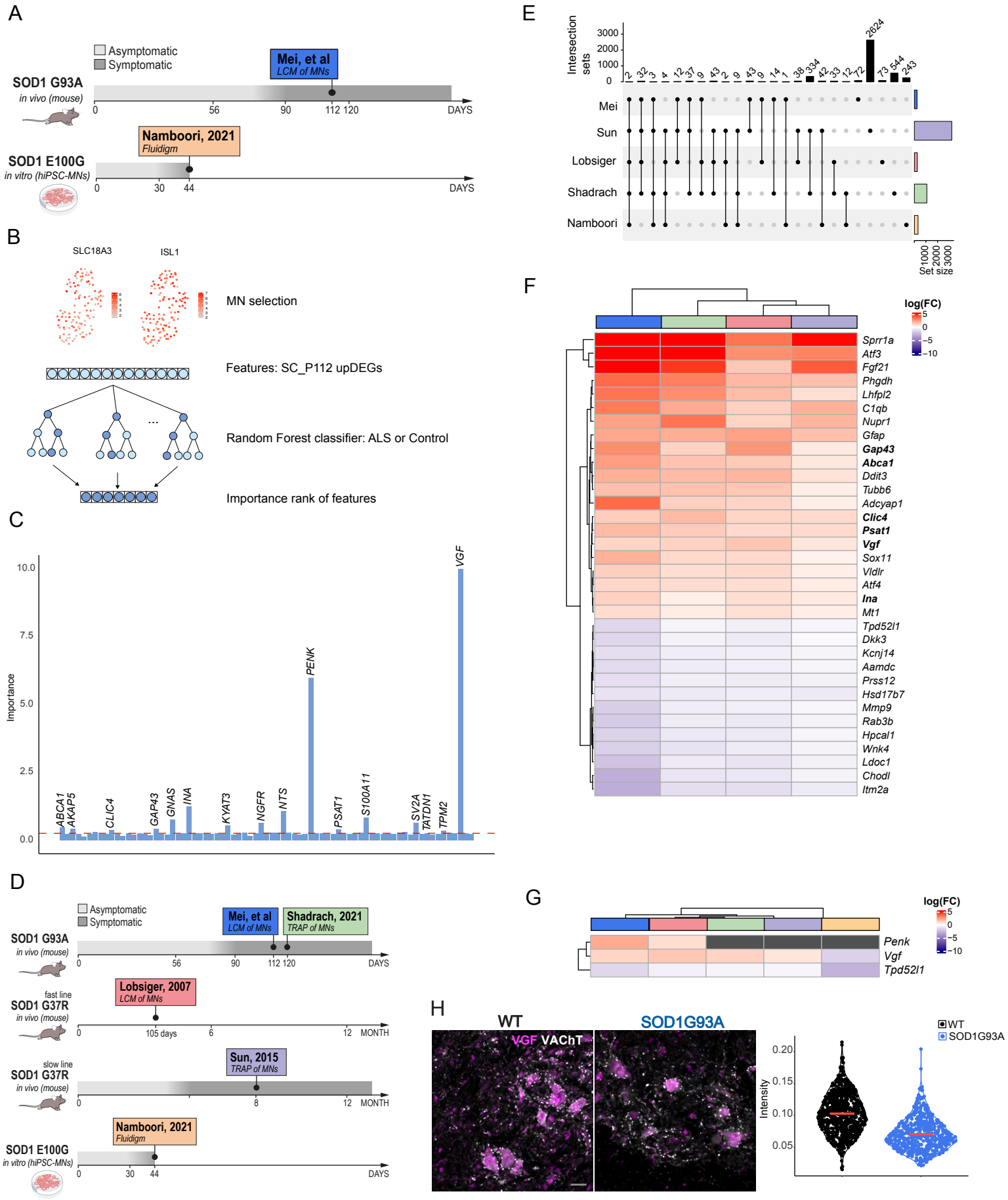
A
Asymptomatic
Symptomatic
SOD1 G93A
in vivo (mouse)
Mei, et al
LCM of MNs
0
56
90
112 120
DAYS
Namboori, 2021
Fluidigm
SOD1 E100G
in vitro (hiPSC-MNs)
0
30
44
DAYS
B
SLC18A3
ISL1
MN selection
Features: SC_P112 upDEGs
Random Forest classifier: ALS or Control
Importance rank of features
C
Importance
10.0
7.5
5.0
2.5
0.0
ABCA1
AKAP5
CLIC4
GAP43
GNAS
INA
KYAT3
NGFR
NTS
PENK
PSAT1
S100A11
SV2A
TATDN1
TPM2
VGF
D
Asymptomatic
Symptomatic
SOD1 G93A
in vivo (mouse)
Mei, et al
LCM of MNs
Shadrach, 2021
TRAP of MNs
0
56
90
112 120
DAYS
fast line
SOD1 G37R
in vivo (mouse)
Lobsiger, 2007
LCM of MNs
0
105 days
6
12
MONTH
slow line
SOD1 G37R
in vivo (mouse)
Sun, 2015
TRAP of MNs
0
8
12
MONTH
SOD1 E100G
in vitro (hiPSC-MNs)
Namboori, 2021
Fluidigm
0
30
44
DAYS
E
Intersection sets
3000
2000
1000
0
2
32
3
4
12
37
9
43
2
9
43
9
14
1
38
334
42
33
12
72
2624
73
544
243
Mei
Sun
Lobsiger
Shadrach
Namboori
1000
2000
3000
Set size
F
log(FC)
5
0
−5
−10
Sprr1a
Atf3
Fgf21
Phgdh
Lhfpl2
C1qb
Nupr1
Gfap
Gap43
Abca1
Ddit3
Tubb6
Adcyap1
Clic4
Psat1
Vgf
Sox11
Vldlr
Atf4
Ina
Mt1
Tpd52l1
Dkk3
Kcnj14
Aamdc
Prss12
Hsd17b7
Mmp9
Rab3b
Hpcal1
Wnk4
Ldoc1
Chodl
Itm2a
G
log(FC)
5
0
−5
−10
Penk
Vgf
Tpd52l1
H
WT
SOD1G93A
VGF VAChT
WT
SOD1G93A
Intensity
0.20
0.15
0.10
0.05

Figure 6