Check for updates

Whole blood transcriptome profile identifies motor neurone disease RNA biomarker signatures

Sulev Kõks[1,2]*, Karin Rallmann[3], Mari Muldmaa[4,5], Jack Price[2], Abigail L. Pfaff[1,2] and Pille Taba[5]

[1]Centre for Molecular Medicine and Innovative Therapeutics, Murdoch University, Perth, WA, Australia, [2]Perron Institute for Neurological and Translational Science, Perth, WA, Australia, [3]Department of Neurology, Tartu University Hospital, Tartu, Estonia, [4]Department of Neurology, North Estonia Medical Center, Tallinn, Estonia, [5]Institute of Clinical Medicine, University Tartu, Tartu, Estonia

Abstract

Blood-based biomarkers for motor neuron disease are needed for better diagnosis, progression prediction, and clinical trial monitoring. We used whole blood-derived total RNA and performed whole transcriptome analysis to compare the gene expression profiles in (motor neurone disease) MND patients to the control subjects. We compared 42 MND patients to 42 aged and sex-matched healthy controls and described the whole transcriptome profile characteristic for MND. In addition to the formal differential analysis, we performed functional annotation of the genomics data and identified the molecular pathways that are differentially regulated in MND patients. We identified 12,972 genes differentially expressed in the blood of MND patients compared to age and sex-matched controls. Functional genomic annotation identified activation of the pathways related to neurodegeneration, RNA transcription, RNA splicing and extracellular matrix reorganisation. Blood-based whole transcriptomic analysis can reliably differentiate MND patients from controls and can provide useful information for the clinical management of the disease and clinical trials.

KEYWORDS

motor neuron disease, amyotrophic lateral sclerosis, RNA-seq, whole transcriptome, gene expression profiling

Impact statement

The present study analysed the gene expression on the whole transcriptome scale in the blood of motor neuron disease (MND) patients. We demonstrated that MND patients have highly specific gene expression patterns or fingerprints, and many genes are differentially expressed in the blood of MND patients. This finding significantly impacts our understanding of the role of the differentially expressed genes in the

pathogenesis of MND. These findings present the utility of RNA-base blood biomarkers for neurological diseases and in precision clinical management.

Introduction

Motor neurone disease (MND) is a group of chronic sporadic and familial disorders characterised by progressive degeneration of motor neurons [1]. The disease is caused by the degeneration of the upper, lower, or both motor neurones. The prognosis of MND depends upon the age at onset and the area of the central nervous system affected [2]. Based on the site of origin and the severity of neurological involvement, four main subtypes of MND have been described: amyotrophic lateral sclerosis (ALS), progressive bulbar palsy (PBP), progressive muscular atrophy (PMA), and primary lateral sclerosis (PLS) [3].

ALS is the most common form of MND. ALS and MND are commonly used interchangeably or as synonyms. ALS is also known as Lou Gehrig's disease or Charcot disease [1]. ALS is an adult-onset, progressive, neurodegenerative disorder involving the large motor neurons of the brain and the spinal cord. It produces a characteristic clinical picture with weakness and wasting of the limbs and bulbar muscles, leading to death from respiratory failure within 5 years.

The degeneration of motor neurons is irreversible, and apparently, it starts many years before the clinical features emerge. Therefore, reliable biomarkers from easily accessible tissues are needed for earlier diagnosis and better prediction of the progression of the disease. The molecular pathology underlying MND relies on genetic variants described in at least 100 different genes to date and on the overlay of the transcriptomic changes [4, 5]. The pathogenesis of the disease involves oxidative stress, inflammation, ER stress with protein aggregation, autophagy and aberrant RNA processing [5, 6].

Familial and sporadic forms of MND can be distinguished based on the evidence of genetic variants and family history [7]. However, only about 20% of MND cases can be explained by known genetic variations [8].

In addition to the well-known genes and their variants, we recently described an unexpectedly large number of exonisation of SINE-VNTR-Alu repeats (SVAs) in the motor cortex [9]. SVAs are known to alter splicing, and several of these elements have been associated with disease through such mechanisms [10, 11]. This indicates the significant role that noncoding or dark genomes can play in the pathogenesis of complex diseases. Moreover, analysis of the whole transcriptome gives an excellent functional opportunity to explore the molecular changes at different stages of diseases, making it a suitable tool for biomarkers [10]. Indeed, transcriptomic analysis can be performed from any biological material, like blood or cerebrospinal fluid and can be used for different conditions [5, 12, 13]. Transcriptomic analysis helps to understand the effect of DNA variants, especially for the splicing-altering variants.

Post-mortem tissue analysis for chronic diseases is always an option to identify molecular patterns in the affected tissues, and this can help to classify the different pathogenic mechanisms [6]. However, using peripheral tissues, like blood, skin, or saliva, allows molecular profiling during the disease's progression and real-life monitoring of pathogenic changes [12, 14, 15]. In the case of MND, several previous studies have been performed to analyse the transcriptomic profile of the blood [16, 17]. In one example, whole blood-derived RNA (PAXgene tubes) was used for microarray analysis; in another, PBMC-derived RNA was used for RNA sequencing. These studies have their limitations. In the case of the microarray analysis, only a certain number of genes that are printed in the microarray can be analysed, and while the number is high (29,830 unique and suitable probes), the

TABLE 1 General characteristics of the study cohort.

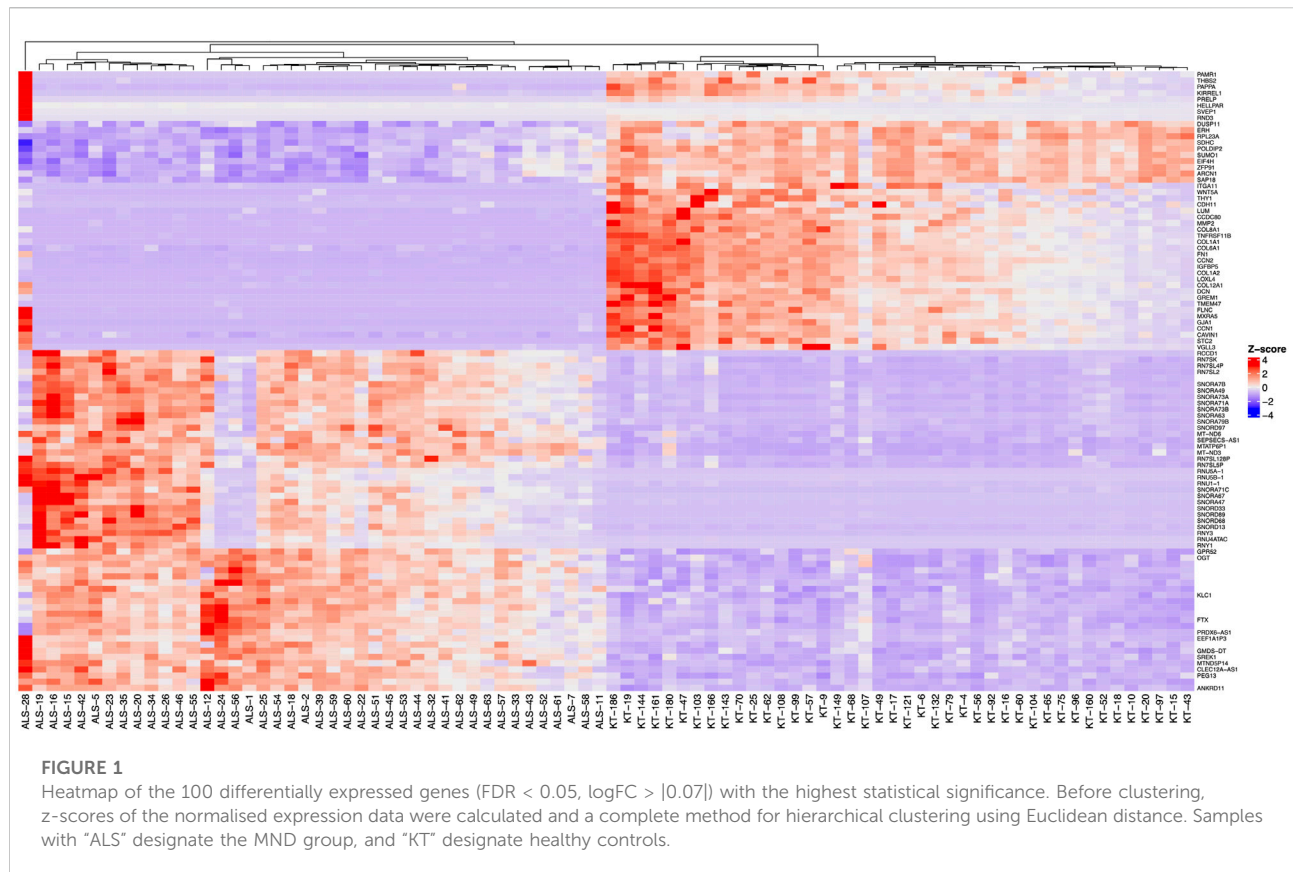| Group	Motor neurone disease (MND)	Healthy controls (HC)
Total, n[a]	42	42
Male, n	13	13
Female, n	29	29
Total mean age, y (sd)	65.6 (9.3)	65.7 (9.4)
Male mean age, y (sd)	64.6 (11.6)	64.3 (11.4)
Female mean age, y (sd)	66.2 (8.4)	66.2 (8.4)
MND duration, m (sd)	18.2 (19.2)	—
Male duration, m (sd)	16.6 (24.5)	—
Female duration, m (sd)	19.0 (16.8)	—

[a]n, number of subjects; y, years; m, months; sd, standard deviation.

TABLE 2 Differentially expressed genes in the blood of MND patients compared to healthy controls. The top 30 genes are shown sorted by the FDR-adjusted *p*-value.

Ensembl ID	logFC	p-adj	Gene name	Gene symbol
ENSG00000202354	6.49	1.25E-136	RNA, Ro60-associated Y3	RNY3
ENSG00000201098	7.26	4.63E-119	RNA, Ro60-associated Y1	RNY1
ENSG00000282885	3.29	3.44E-111	novel transcript	lnc-NEMF-1
ENSG00000091986	−6.81	4.27E-101	coiled-coil domain containing 80	CCDC80
ENSG00000011465	−7.48	9.56E-99	decorin	DCN
ENSG00000118523	−4.88	5.04E-97	cellular communication network factor 2	CCN2
ENSG00000164692	−7.26	1.58E-95	collagen type I alpha 2 chain	COL1A2
ENSG00000108821	−6.43	1.04E-93	collagen type I alpha 1 chain	COL1A1
ENSG00000128591	−5.83	8.20E-93	filamin C	FLNC
ENSG00000138131	−5.55	1.70E-87	lysyl oxidase like 4	LOXL4
ENSG00000113739	−7.73	2.01E-86	stanniocalcin 2	STC2
ENSG00000199568	9.05	1.07E-80	RNA, U5A small nuclear 1	RNU5A-1
ENSG00000248527	3.20	6.12E-80	MT-ATP6 pseudogene 1	MTATP6P1
ENSG00000087245	−6.83	6.87E-80	matrix metallopeptidase 2	MMP2
ENSG00000186340	−7.55	2.94E-79	thrombospondin 2	THBS2
ENSG00000115414	−5.71	1.45E-78	fibronectin 1	FN1
ENSG00000150459	−0.79	3.01E-77	Sin3A associated protein 18	SAP18
ENSG00000212283	5.02	4.23E-77	small nucleolar RNA, C/D box 89	SNORD89
ENSG00000111799	−6.88	1.03E-75	collagen type XII alpha 1 chain	COL12A1
ENSG00000199631	7.52	2.52E-73	small nucleolar RNA, C/D box 33	SNORD33
ENSG00000144810	−6.44	3.84E-73	collagen type VIII alpha 1 chain	COL8A1
ENSG00000164761	−6.80	1.42E-72	TNF receptor superfamily member 11b	TNFRSF11B
ENSG00000115963	−6.65	2.38E-70	Rho family GTPase 3	RND3
ENSG00000115461	−7.69	9.88E-70	insulin like growth factor binding protein 5	IGFBP5
ENSG00000126214	0.95	1.05E-63	kinesin light chain 1	KLC1
ENSG00000186660	−0.54	1.50E-63	ZFP91 zinc finger protein, E3 ubiquitin ligase	ZFP91
ENSG00000142156	−2.90	1.73E-62	collagen type VI alpha 1 chain	COL6A1
ENSG00000238961	5.94	3.17E-62	small nucleolar RNA, H/ACA box 47	SNORA47
ENSG00000166923	−7.07	2.09E-61	gremlin 1, DAN family BMP antagonist	GREM1
ENSG00000281501	1.98	5.08E-61	SEPSECS antisense RNA 1	SEPSECS-AS1

whole transcriptome sequencing gives information entire transcriptome (60,230 elements) [17]. Moreover, RNA-seq has a better dynamic range in detecting gene expression therefore the power to detect differential expression is better. PBMC-derived samples only include monocytes and do not contain neutrophils, basophils, and eosinophils. While basophils and eosinophils are only a small subset of all immune cells (0–2% and 1–7%, respectively), neutrophils make up a majority of circulating nucleated blood cells (45–75%) [18]. Therefore, analysing PBMC samples will give only partial information about the RNA changes in the blood and this has been shown in many studies [18–20]. The present study aimed to perform whole transcriptome analysis from the whole-blood (Tempus tubes) derived RNA and to identify the

FIGURE 1
Heatmap of the 100 differentially expressed genes (FDR < 0.05, logFC > |0.07|) with the highest statistical significance. Before clustering, z-scores of the normalised expression data were calculated and a complete method for hierarchical clustering using Euclidean distance. Samples with "ALS" designate the MND group, and "KT" designate healthy controls.

whole blood transcriptomic profile by comparing MND patients to the age and sex-matched healthy controls.

Materials and methods

Study cohort

Between 2013 and 2018, a total of 84 participants (42 MND patients and 42 healthy control patients without any chronic diseases) were enrolled in the study and signed written informed consent. Inclusion criteria for MND patients were the diagnosis of probable or definitive MND based on El Escorial Criteria and the absence of a positive family history.

For the healthy controls, we excluded patients with any chronic diseases, especially any neurologic, rheumatological, haematological, or oncological conditions. In addition, treatment with biologics or chemotherapy was also excluded. A white blood cell (WBC) count and C-reactive protein (CRP) were measured in every health control to exclude any underlying inflammatory condition.

The blood samples were collected into Tempus Blood RNA tubes and stored according to the manufacturer's instructions. The research was conducted with the approval of the University of Tartu Research Ethics Committee, and all participants provided written informed consent. The comprehensive
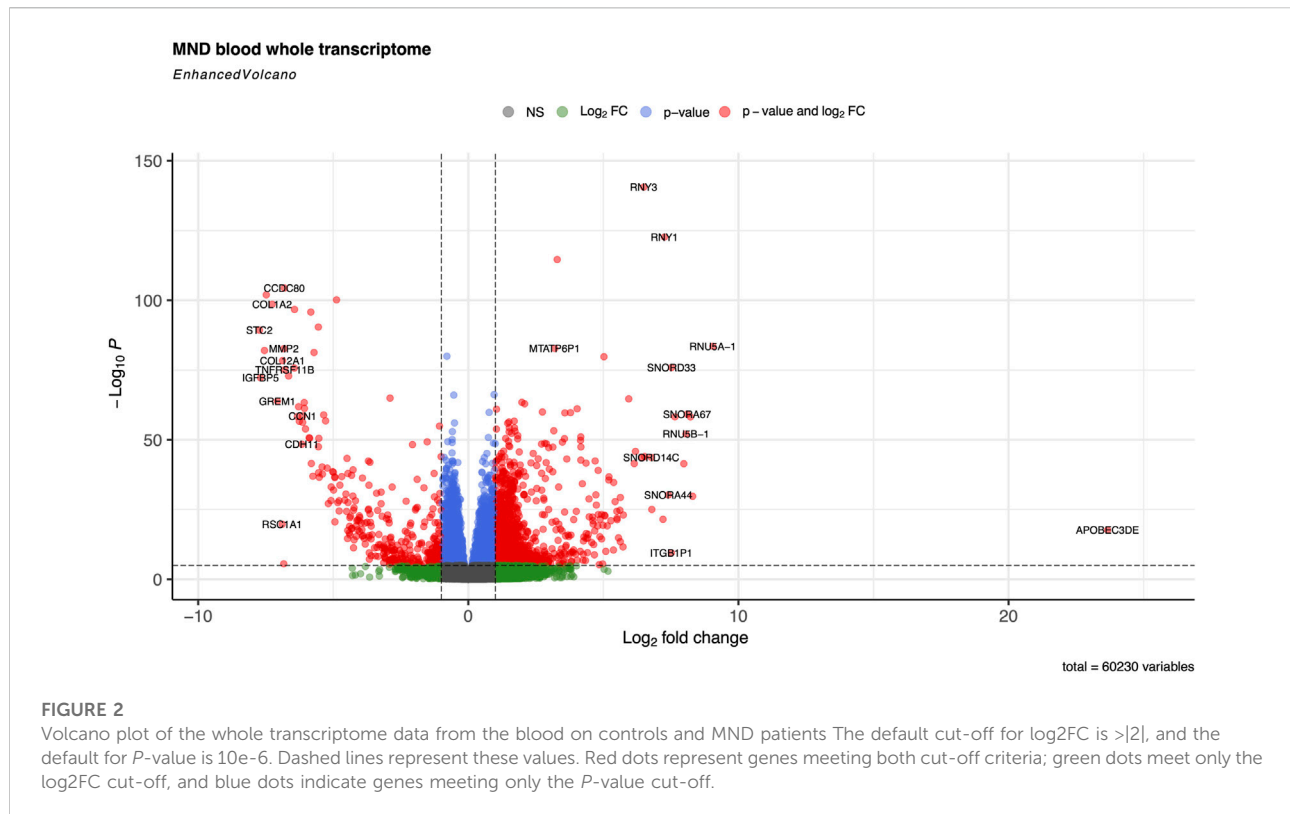
patient selection process leveraged hospital records, neurologist consultations, and the Estonian Health Insurance Fund's national health data repository.

The whole blood was collected from 42 MND patients and 42 healthy controls using Tempus Blood RNA collection tubes (Thermo Fisher Scientific). Neurologists recruited MND patients, and the subtype of the MND was confirmed. Healthy controls were recruited among the visitors referred to the blood analysis who did not have chronic diseases. The control samples were ideal controls without any neurological condition or major chronic illness and were age- and sex-matched to the MND group (complete information is given in Supplementary Table S1).

Whole transcriptome analysis and functional annotation

The RNA was isolated from whole blood using a Tempus Spin Isolation Kit (Thermo Fisher Scientific). After initial quality control and quantification (A260/280 ratio, RIN number). RNA was used for the total RNA sequencing necessary for the whole transcriptome analysis.

Total RNA sequencing was performed in all 84 samples at the Genomics Core Facility at Murdoch University, Perth, WA. Illumina paired-end 2 × 100bp read length using NovaSeq 6000.

FIGURE 2
Volcano plot of the whole transcriptome data from the blood on controls and MND patients The default cut-off for log2FC is >|2|, and the default for *P*-value is 10e-6. Dashed lines represent these values. Red dots represent genes meeting both cut-off criteria; green dots meet only the log2FC cut-off, and blue dots indicate genes meeting only the *P*-value cut-off.

The NovaSeq Control Software v1.7.5 and Real-Time Analysis (RTA) v3.4.4 performed real-time image analysis. RTA performs real-time base calling on the NovaSeq instrument computer. The Illumina DRAGEN BCL Convert 07.021.624.3.10.8 pipeline generated the sequence data. The FASTQ files were analysed using salmon 1.10.3 by using the reference genome GRCh38 [21]. Salmon counts were imported to the R studio using the *tximeta* package [22]. Differential whole transcriptome analysis was performed with the *DESeq2* package [23]. No fold-change filtering was initially applied, but the False Discovery Rate (FDR) was set at 0.05 to adjust for multiple testing, and this corresponds to the 1.05 fold change threshold in our experiment.

The functional annotation of the differential gene expression was performed with the packages *ReactomePA*, *clusterProfiler* and *DOSE* [24–26]. Principal component analysis was performed by using *pcaExplorer* and *factoextra* packages. The heatmap clustering was performed with the *ComplexHeatmap* package based on the z-scores of the normalised expression data and using Euclidean distance for complete linkage agglomerative clustering.

Pair-wise analysis

To perform a pair-wise analysis of individual genes between MND and healthy controls, we applied the two-tailed Wilcoxon rank-sum test implemented in the function compare_means() of

the package *ggpubr* [27]. We generated a list of all known MND genes using the OMIM catalogue and identified 97 genes that are directly connected to the MND or its subtypes. This list extracted normalised counts from the salmon quant files and made boxplots with pairwise comparisons. Plots were generated using *ggplot2* version 3.5.1 and *ggpubr* version 0.6.0 packages. Statistical analysis was performed with R software version 4.4.0 and RStudio Version 2023.06.0 + 421.
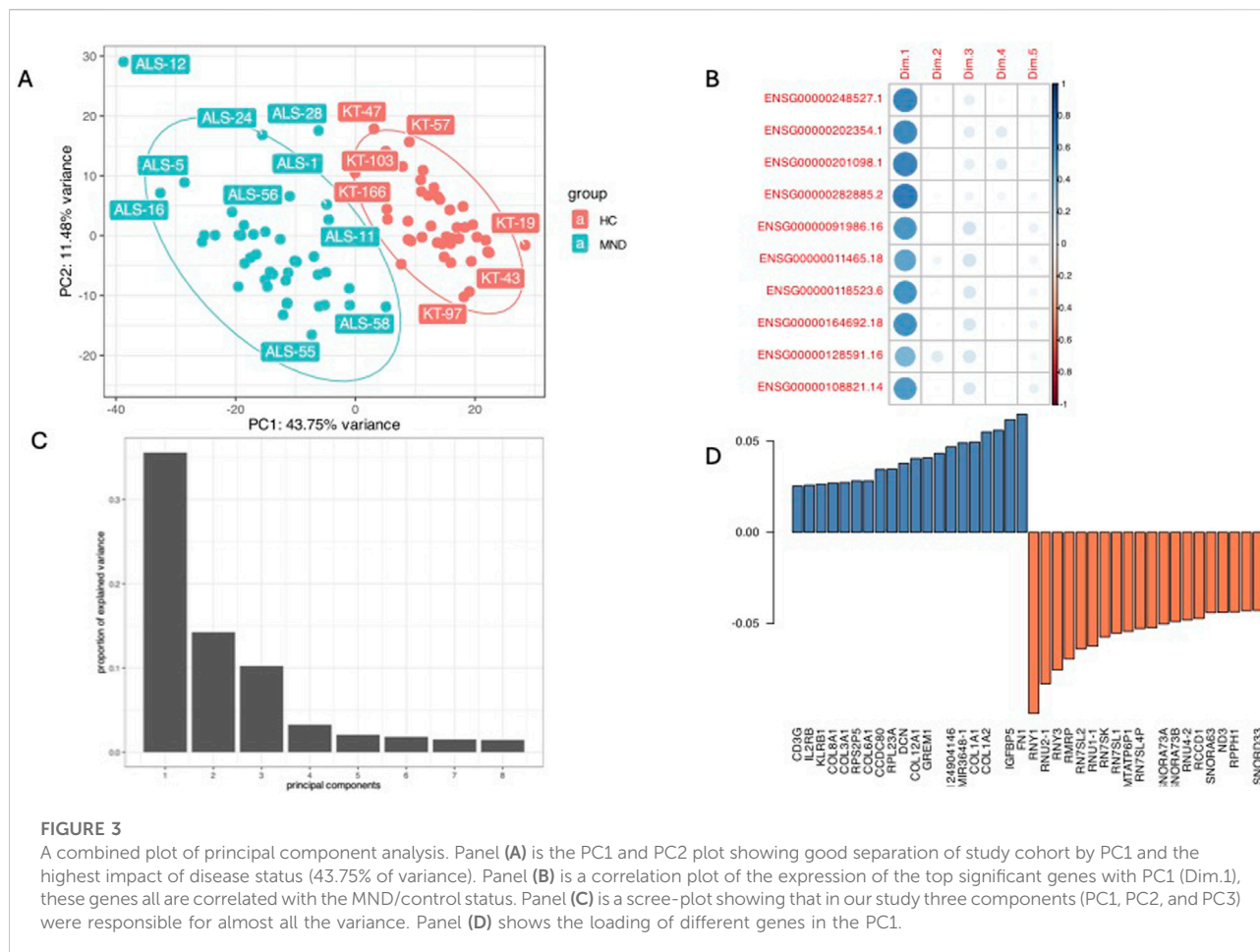
Results

Description of the study cohort

The general characteristics of the population are reported in Table 1. The median age was 65.6 (standard deviation 9.3) years, and most subjects were female (69%). No patient reported a positive family history of MND; therefore, all the participants had sporadic forms, and all patients received standard MND therapy with riluzole. The most frequent clinical subtype was the classic ALS (86%). Spinal symptoms were present the most commonly (60%).

Whole blood RNA sequencing

RNA sequencing resulted in at least 50 million paired 150 bp reads per sample, and all reads had Phred score higher than 30.

FIGURE 3
A combined plot of principal component analysis. Panel **(A)** is the PC1 and PC2 plot showing good separation of study cohort by PC1 and the highest impact of disease status (43.75% of variance). Panel **(B)** is a correlation plot of the expression of the top significant genes with PC1 (Dim.1), these genes all are correlated with the MND/control status. Panel **(C)** is a scree-plot showing that in our study three components (PC1, PC2, and PC3) were responsible for almost all the variance. Panel **(D)** shows the loading of different genes in the PC1.

Salmon was used to quantify transcript abundances from fastq files. *Tximeta* was used to import the resulting quant files, and gene-level summarisation was used for the *DESeq2* workflow. Healthy controls were compared to the MND RNA-seq results, and we identified 12,972 genes differentially expressed (FDR < 0.05) in the blood of MND patients. The top 30 differentially regulated genes are shown in Table 2. Out of these 12,972 genes, 8,008 were upregulated, and 4,964 were down-regulated (Supplementary Table S2, sheet 1). A heat map with all 12,972 genes is shown as Supplementary Figure S1, and it shows a clear separation of MND patients from the healthy controls. A smaller heatmap with the top 100 genes is shown in Figure 1, and a volcano plot is shown in Figure 2. The heatmap with 100 genes shows a consistent and clear separation of the MND from the healthy controls. This remarkable finding shows that a disease highly specific to the central nervous system can be differentiated from controls by the blood transcriptome profile.

When we used the FDR 0.05 filtering threshold, we detected the genes Log2 FC 0.07 threshold, which transforms to an expression difference of 1.05-fold change ($2^{0.07}$). We then applied an additional fold change threshold to filter the dataset further. When we applied FC threshold of 1.1

(log2 FC 0.13), we got 12,839 differentially expressed genes (DEGs, Supplementary Table S2, sheet 2). With the FC threshold of 1.5 (log2 FC 0.59), we got 6,403 DEGs (Supplementary Table S2, sheet 3), and finally, applying the threshold of FC 2.0 (log2 FC 1.0), we got 3,286 DEGs (Supplementary Table S2, sheet 4).

The principal component analysis identified that disease status, PC1 was responsible for 43.75% of the variance and gene expression profiles clearly separated MND patients from healthy controls (Figure 3A). The genes with the highest differential expression (the lowest FDR values) had a very high correlation with the PC1 (Figure 3B, "Dim.1" is PC1) and the scree plot (Figure 3C) verified that most of the variation in our study cohort is explained by three principal components, PC1, PC2 and PC3. PC1 is disease status, and we were not able to identify the essence of the PC2 and PC3. These are neither the sex (Supplementary Figure S2, Panels A) nor age (Supplementary Figure S2, Panels B and C) of the patients, nor the type of the disease (Supplementary Figure S2, Panels D and E). It could be that PC2 and PC3 are some other factors reflecting the heterogeneity of the pathophysiology of the MND.
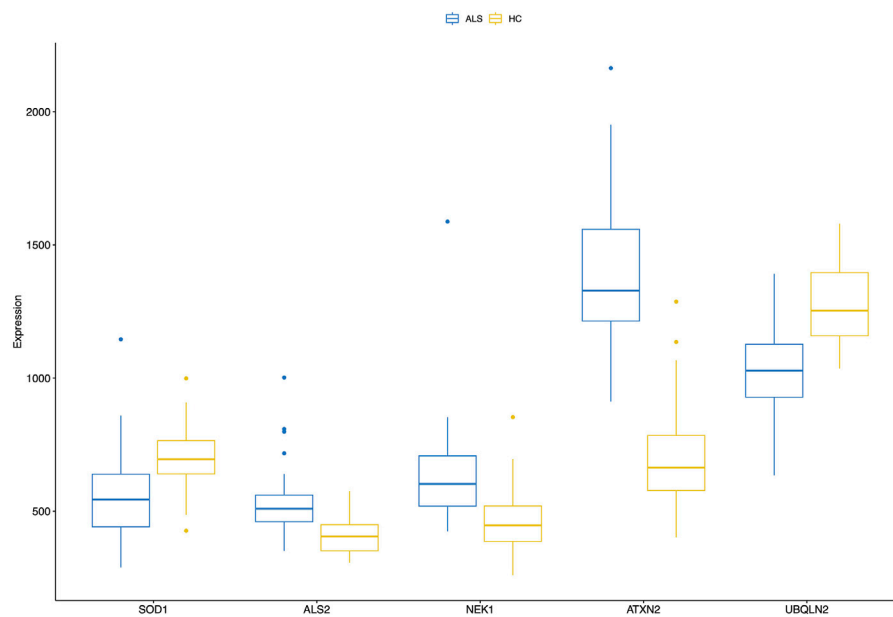
FIGURE 4
A combined boxplot of five MND-related genes and their expression levels in the blood of MND patients and controls gives comparative blood expression levels for these selected genes. Pairwise statistical comparisons are shown in Figure 5 and in Supplementary Figure S3.

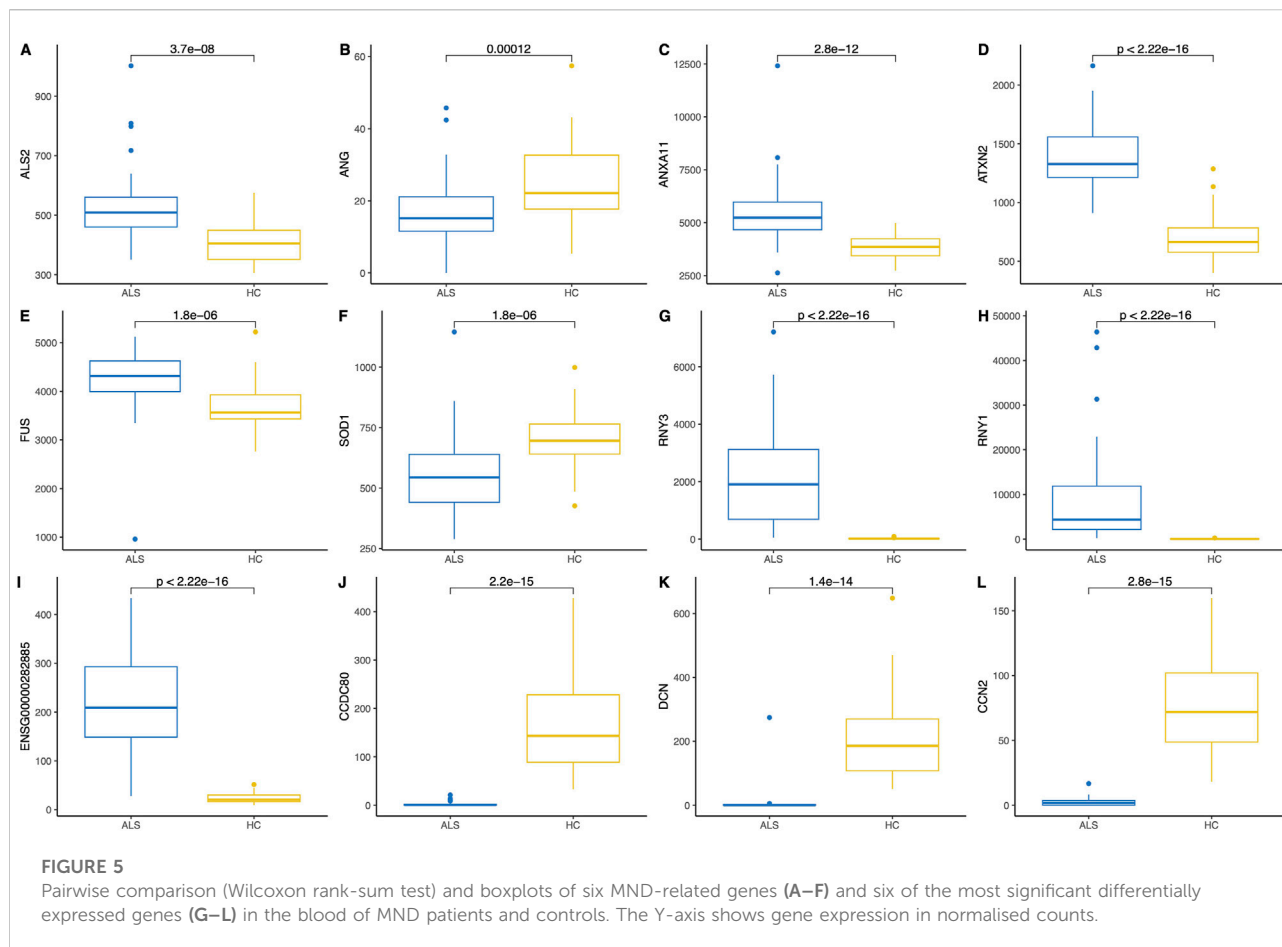## Pairwise analysis of known MND genes

In addition to the whole transcriptome analysis, we performed a pairwise (MND versus healthy controls) study of 97 known MND genes (a list of the genes is provided in Supplementary Table S3) and 30 top-regulated genes from the DESeq2 analysis. All results are shown in Supplementary Figure S3, and partial results are in Figures 4, 5. Interestingly, some MND-related genes are upregulated (*ALS2*, *NEK1*, *ATXN2*), while others are downregulated (*SOD1*, *UBQLN2* aka *ALS15*) in patients. In addition, *FUS* and *ANXA11* were upregulated, and *ANG* was downregulated in patients (Figures 5A–F). Moreover, the DESeq2 top genes *RNY3*, *RNY1*, and *ENSG0000282885* were highly upregulated in patients with almost no expression in control subjects (Figures 5G–I). At the same time, other DESeq2 top genes, *CCDC80*, *DCN*, and *CCN2*, were highly expressed in controls, and their expression was almost missing in patients' blood (Figures 5J–L). These examples indicate that there are many high fold-change difference genes with almost no expression in one group and very high expression in another, and these genes have very high potential to be a transcriptional biomarker for the MND.

The pairwise analysis of all 97 MND genes indicated that some well-known MND genes weren't differentially expressed in the blood (boxplots are in Supplementary Figure S3). Out of all 97 genes, 38 (39%) of them *AMFR, AR, ATX3, BICD2, C9orf72, CHRNA3, DAO, DCTN1, DNAJC7, ERBB4, HNRNPA2B1, IGFALS, KIF5A, LGALSL, LRP12, MAPT, MOBP, NEFH, OPTN, PAH, PON1, PON2, PON3, PRPH,* *PSEN1, SARM1, SCYL1, SETX, SLC1A2, SLC52A3, SMN1, SMN2, SQSTM1, TARDBP, TRPM7, TUBA4A, VRK1, VSX1,* were not differentially expressed between patients and controls. Fourteen genes of these 38 genes were not expressed in blood. Most of these genes that were not differentially expressed had excellent expression levels in the blood. *AMFR* has an expression level of 1,800 normalised counts, *C9orf72* has 1,500 normalised counts, *PSEN1* has an expression at 2,500 normalised counts, *TARDBP* has an average gene expression of 1,600 normalised counts, *SQSTM1* has an expression level of 3,100 normalised counts. Therefore, all these genes are highly expressed in the whole blood, but their expression level is not dependent on the disease status.

Functional annotation of differentially expressed genes

Functional annotation of differentially expressed genes indicated statistically significant activation of several human disease pathways (Table 3, full version provided in Supplementary Table S4). Remarkably, three neurodegenerative diseases were at the top of the table of the KEGG pathways: Parkinson's disease, prion disease, and amyotrophic lateral sclerosis (Figure 6). In addition, several pathways involved in the pathogenesis of neurodegeneration were also activated. These included protein processing in the endoplasmic reticulum, proteasome, lysosome and ubiquitin-mediated proteolysis.

FIGURE 5
Pairwise comparison (Wilcoxon rank-sum test) and boxplots of six MND-related genes **(A–F)** and six of the most significant differentially expressed genes **(G–L)** in the blood of MND patients and controls. The Y-axis shows gene expression in normalised counts.

Reactome and GSEA analyses use more canonical pathways (Supplementary Tables S5, S6). Reactome identified statistically significant enrichment of the mRNA splicing and transcription-related pathways in combination with cellular energetics pathways (mitochondria and respiratory electron transport) to be affected (Figure 7). GSEA analysis (Figure 8) identified statistically significant enrichment of sensory perception, olfactory signalling and many pathways related to the extracellular matrix reorganisation (collagen degradation, elastic fibre formation, assembly of collagen fibres).

In summary, KEGG pathway analysis found statistically significant activation of the ALS pathway together with other neurodegeneration pathways. The findings from Reactome and GSEA added more details to the KEGG finding and identified several cellular pathways that can give a mechanistic understanding of the pathogenesis of MND.

Discussion

The current study presented a whole transcriptome analysis of the whole blood RNA from MND patients compared to age and sex-matched healthy controls

(Figure 9). As a main finding, we identified 12,972 genes differentially expressed; 8,008 were upregulated, and 4,964 were downregulated in the blood of MND patients. Most remarkably, the heatmap based on these 12,972 genes was highly specific and separated MND from healthy controls. Therefore, we can conclude that the identified differentially expressed genes are specific for the MND status. This doesn't mean that all of these genes are directly related to the pathogenesis of MND but instead reflects the complexity of the disease, where pathogenic changes are mixed with compensatory changes. However, this still shows that MND, while a CNS-specific disease, has remarkable changes in the blood transcriptomics, and blood could be a perfect source for the diagnostic biomarkers for MND.

The number of differentially expressed genes seems to be unreasonably high, but van Rheenen et al., used Illumina bead chips with only 29,830 unique and suitable probes, and they also identified 7,038 genes to be differentially expressed [17]. This number is very close to the one that we identified if we take into account that in our study, we used RNA-seq that analysed the expression of 60,230 genes, and our sample is perfect sex and age-matched, which means more power. In addition, in our own previous study, we identified

TABLE 3 KEGG pathways that are enriched in the blood transcriptome of MND patients.

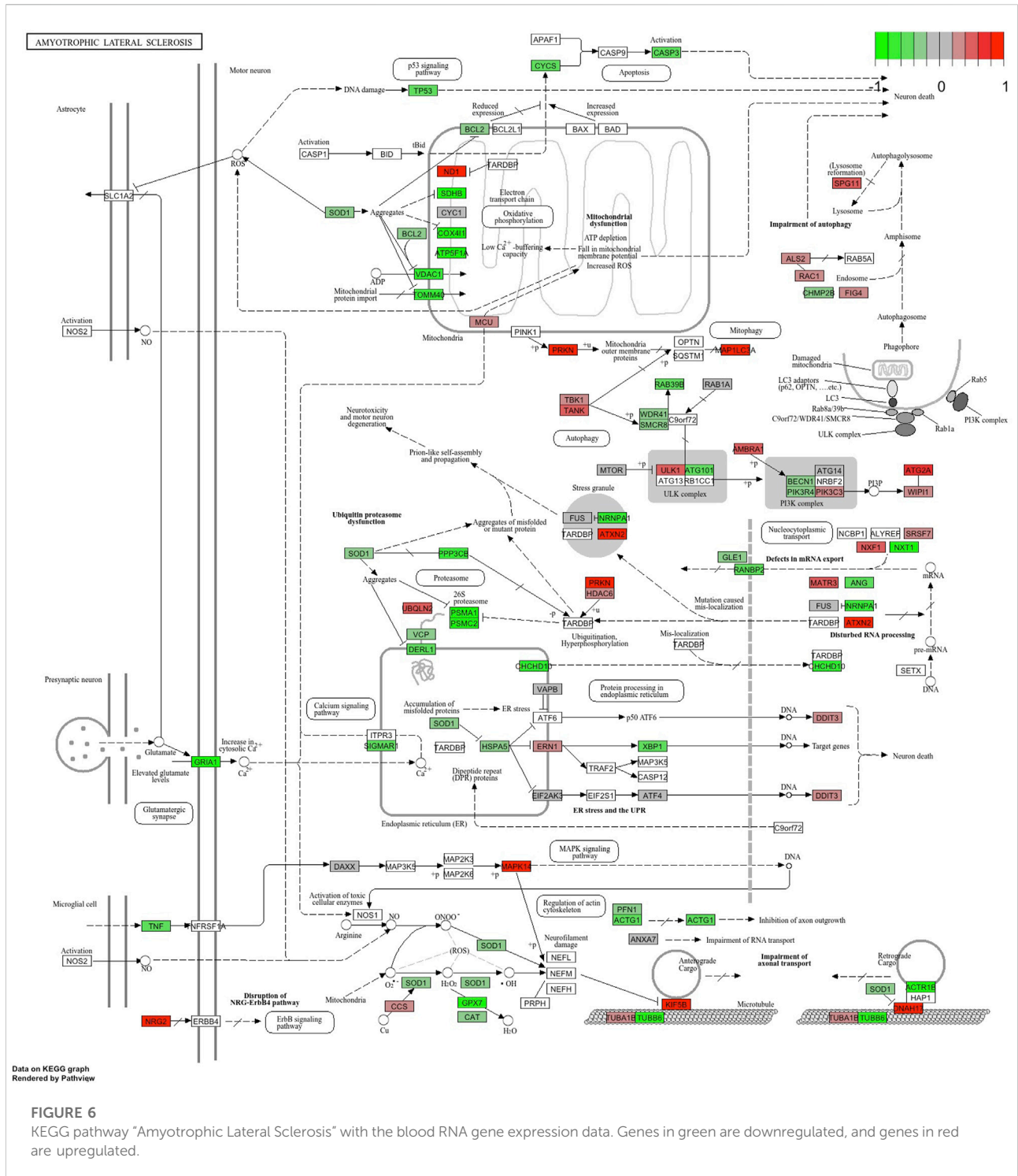| Category	Subcategory	ID	Description	Gene ratio	Bg ratio	P-adjusted
Human Diseases	Neurodegenerative disease	hsa05012	Parkinson disease	165/3,781	271/8,843	2.36E-07
Human Diseases	Cardiovascular disease	hsa05415	Diabetic cardiomyopathy	130/3,781	205/8,843	2.36E-07
Human Diseases	Neurodegenerative disease	hsa05020	Prion disease	167/3,781	278/8,843	3.14E-07
Human Diseases	Neurodegenerative disease	hsa05014	Amyotrophic lateral sclerosis	212/3,781	371/8,843	5.97E-07
Metabolism	Energy metabolism	hsa00190	Oxidative phosphorylation	92/3,781	138/8,843	5.97E-07
Organismal Systems	Environmental adaptation	hsa04714	Thermogenesis	143/3,781	235/8,843	5.97E-07
Genetic Information Processing	Folding, sorting and degradation	hsa04141	Protein processing in endoplasmic reticulum	109/3,781	170/8,843	5.97E-07
Human Diseases	Neurodegenerative disease	hsa05022	Pathways of neurodegeneration - multiple diseases	265/3,781	483/8,843	1.07E-06
Human Diseases	Cancer: overview	hsa05208	Chemical carcinogenesis - reactive oxygen species	136/3,781	226/8,843	2.72E-06
Human Diseases	Neurodegenerative disease	hsa05010	Alzheimer disease	217/3,781	391/8,843	5.09E-06
Human Diseases	Neurodegenerative disease	hsa05016	Huntington disease	177/3,781	311/8,843	7.00E-06
Genetic Information Processing	Translation	hsa03013	Nucleocytoplasmic transport	72/3,781	108/8,843	1.14E-05
Genetic Information Processing	Folding, sorting and degradation	hsa03050	Proteasome	36/3,781	46/8,843	2.46E-05
Cellular Processes	Transport and catabolism	hsa04142	Lysosome	82/3,781	132/8,843	0.0001
Genetic Information Processing	Folding, sorting and degradation	hsa04120	Ubiquitin mediated proteolysis	87/3,781	142/8,843	0.0001
Genetic Information Processing	Translation	hsa03010	Ribosome	102/3,781	172/8,843	0.0002
Genetic Information Processing	Chromosome	hsa03082	ATP-dependent chromatin remodeling	73/3,781	117/8,843	0.0003
Cellular Processes	Cell growth and death	hsa04110	Cell cycle	94/3,781	158/8,843	0.0003
Genetic Information Processing	Replication and repair	hsa03030	DNA replication	28/3,781	36/8,843	0.0004
Human Diseases	Infectious disease: bacterial	hsa05132	*Salmonella* infection	138/3,781	251/8,843	0.0009
Metabolism	Glycan biosynthesis and metabolism	hsa00510	N-Glycan biosynthesis	37/3,781	53/8,843	0.0010
Human Disease	Neurodegenerative disease	hsa05017	Spinocerebellar ataxia	84/3,781	144/8,843	0.0016

Bg, background.

4,824 differentially expressed genes in the CSF of MND patients [5]. Therefore, the number of differentially expressed genes between MND patients and healthy controls seems to be high, but also other studies have found a similarly high number of differentially expressed genes.

In addition, the number of differentially expressed genes remains high even after applying different filtering criteria. While we initially did not use any specific fold-change filtering, the statistically significant FD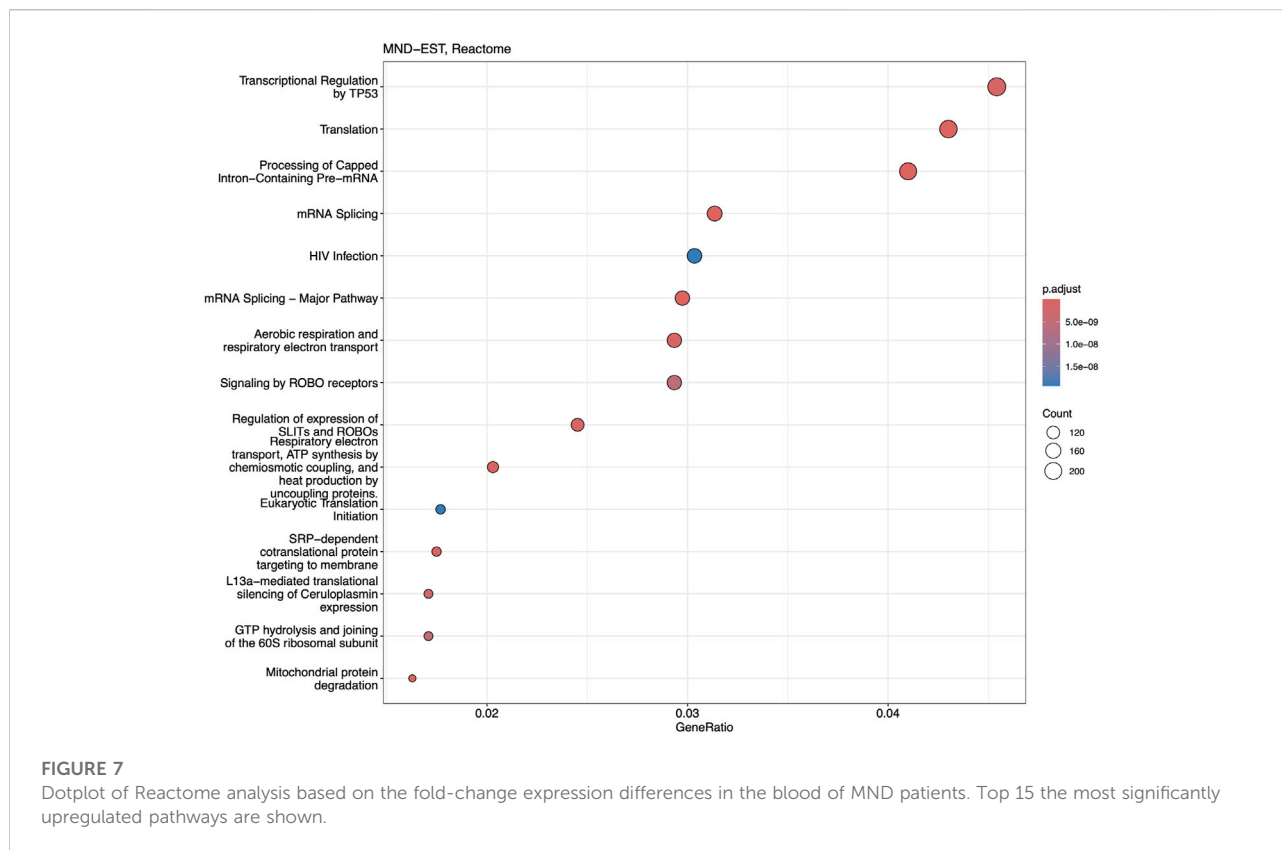R only detected genes with at least a 1.05 fold change difference. When we applied more stringent FC filtering thresholds, the number of differentially expressed genes reduced, but it was still remarkable, with 6,403 genes for FC 1.5 and 3,286 genes with the threshold of FC 2.0. This indicates a robustly specific gene expression profile in the blood of MND patients, making it a reliable source for potential RNA-based biomarkers.

The genes that we identified differentially expressed correlate quite well with the results of the previously published similar studies. We identified all the genes found

FIGURE 6
KEGG pathway "Amyotrophic Lateral Sclerosis" with the blood RNA gene expression data. Genes in green are downregulated, and genes in red are upregulated.

in the paper by Garau et al, Table 5 [28]. In addition, we also compared our genes to the study of van Rheenen et al and found that many genes overlapped between these studies [17]. Therefore, our results are generally in very good concordance with previously published studies.

Not all MND-specific genes were differentially expressed. *C9orf72* is a gene with the highest genetic impact in MND, but it was not differentially expressed. *C9orf72* is highly expressed in the blood with an average normalised count of 1,500. Therefore, the low expression level cannot explain the lack
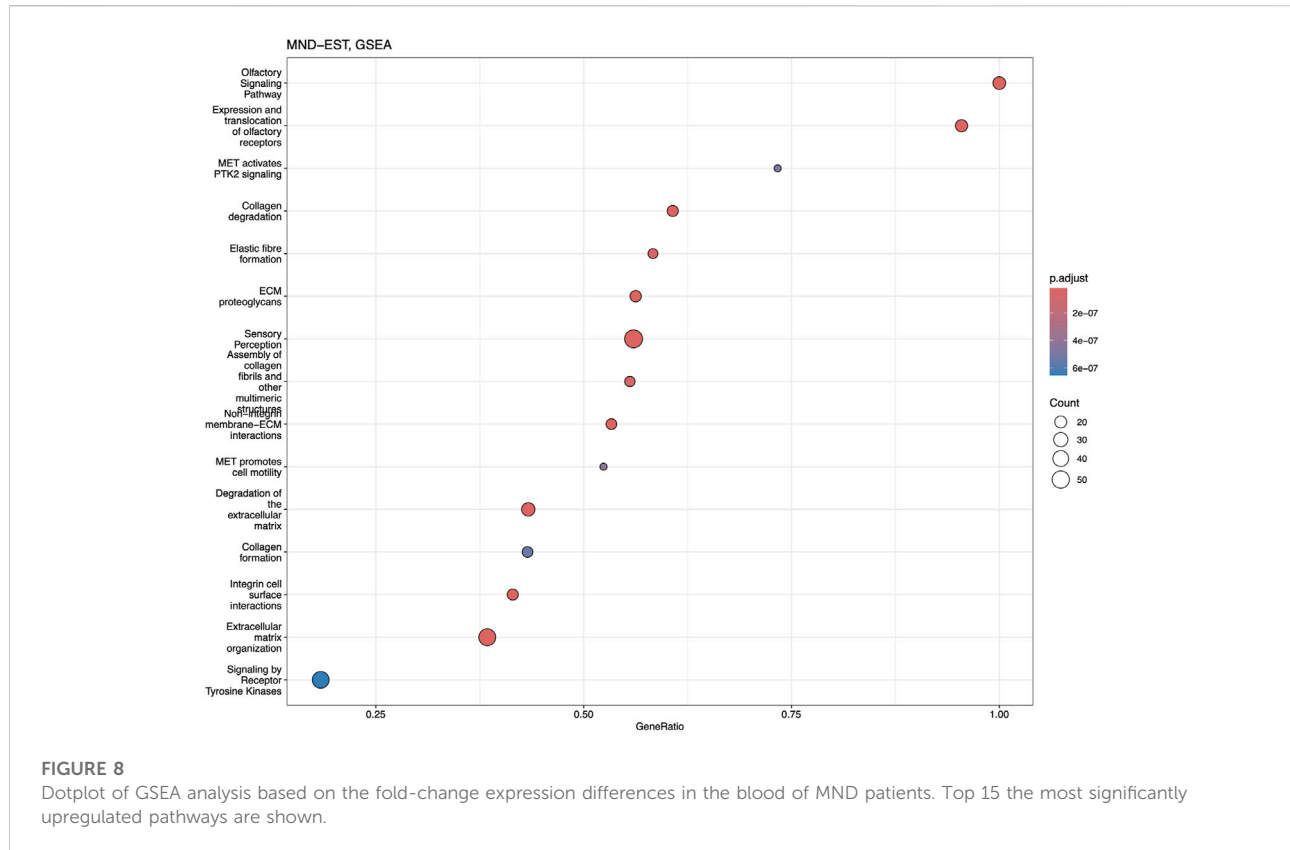
FIGURE 7
Dotplot of Reactome analysis based on the fold-change expression differences in the blood of MND patients. Top 15 the most significantly upregulated pathways are shown.

of significant differences. A similar observation is true for the *SQSTM1*, *TARDBP*, *OPTN* and *PSEN1*, all genes with high expression in the blood, but no difference in expression between MND and controls (Supplementary Figure S2). It is hard to understand why these genes did not show differential expression, but these genes have a mutation-specific effect, and in our cohort, we may not have mutations in these genes. This might be unlikely, as we have identified pathogenic repeat polymorphism for *C9orf72* in one patient who has 1,000 repeats with a length of over 6,000 bp.

We saw significant differences in many MND-related genes. For instance, *SOD1* was downregulated in MND patients. Similarly, *ANG* and *ACSL5* were significantly downregulated in MND patients compared to controls. It is somewhat surprising that *SOD1* is downregulated in MND patients as it is also assumed to form aggregates in sporadic patients [29–31]. At the same time, we couldn't find a significant difference for the *OPTN* gene, another gene that has clear implications in MND pathology and had a very high expression level in blood. It is remarkable that while its aggregates are common for familial and sporadic MND forms, we could not detect significant differences in the expression of *OPTN* [32].

Our study is certainly not the first to analyse MND patients' transcriptomes. One study analysed gene chips

from whole blood RNA, finding 2,943 genes differentially expressed [17]. These authors did not find *SOD1, C9orf72, SQSTM1, TARDBP,* or *OPTN* to be differentially expressed; this study got similar results to ours. Other published studies have used selected cell fractions, like PBMCs or lymphoblastoid cells [16, 28, 33, 34]. The cell fractionation studies identified a much smaller number of differentially expressed genes, and their results are difficult to compare to our results as the approaches are quite different. However, one recent study used a machine learning approach to compare brain and blood transcriptomic data and identified three distinct clusters of the MND subtypes with potentially different pathological mechanisms [6]. These three pathogenic subtypes didn't describe any particular MND mutation but rather the biological pathways that involved particular differentially expressed genes. The present study is based on blood transcriptome, and we have identified similar differentially expressed genes. While we couldn't identify three distinctive subtypes, the heatmap of the 12,972 differentially expressed genes separated MND patients from controls. Moreover, for MND patients, we saw at least two clusters with specific gene expression profiles. Therefore, our study results seem to match the results of the study by Marriott et al [6]. The main finding is that gene expression profiles and RNA analysis could be

FIGURE 8
Dotplot of GSEA analysis based on the fold-change expression differences in the blood of MND patients. Top 15 the most significantly upregulated pathways are shown.

used as a source for biomarkers and can have clinical utility in differentiating patients with distinctive pathogenetic mechanisms.

We identified that the most up-regulated gene, with logFC 23, in MND blood is the *APOBEC3DE* gene (Volcano plot in Figure 2). *APOBEC3DE* is located at 22q13.1 and is a cytidine deaminase gene family member. This gene is one of the APOBEC cluster family on chromosome 22 [35, 36]. APOBEC proteins are part of innate immunity, and they inhibit retroviruses by deaminating cytosine residues in retroviral cDNA [37]. Interestingly, *APOBEC3DE* also inhibits retrotransposition of the long interspersed element-1 (LINE-1) by interacting with ORF1p, a protein encoded by LINE-1 [38]. LINE-1 has been implicated in the pathogenesis of MND, and therefore, *APOBEC3DE* finding seems very relevant as they suppress LINE1 activity [39]. In addition, *APOBEC* proteins can induce somatic mutations into genomic DNA and promote the development of different diseases [40]. *APOBEC* proteins are also involved in the clearance of foreign DNA from human cells, implicating their role in the cellular defence system against mutations that make them very plausible in connection with the MND [41, 42]. Loss of the nuclear *TDP-43* due to the cytoplasmic aggregation of the *TDP-43* is associated with decondensation of the chromatin around LINE1 elements and increased activation or LINE1 with their retrotransposition. Upregulation of the *APOBEC3DE* might be an endogenous

defence mechanism as it is a part of the innate response to retroviral activation [43].

Many differentially expressed genes are involved in splicing and RNA processing: *RNU5A-1*, *RNU1-1*, *RNY3*, and *RNY1*, to name some. Interestingly, these RNA synthesis and splicing-related genes are all upregulated in MND samples and not expressed in the blood of control samples at all. These are genes that have a high potential to become a blood biomarker for MND or help to predict the progression of the disease. While it is not clear how these genes participate in the pathogenesis of MND, splicing mutations and genes participating in splicing involvement in MND have been shown in many previous studies [44–46]. The results from blood transcriptomics were very uniform and showed the upregulation of several genes related to RNA synthesis and splicing, as also indicated in Figure 6.

The function of downregulated genes is more diverse, with possible common denominators being the extracellular matrix (ECM) organisation and remodelling (Figure 7). Reduced expression of *CCDC80*, *COL1A1*, *COL1A2*, *MMP2*, and *TNFRSF11B* indicates the ECM reorganisation also found in GSEA enrichment analysis (Figure 2). The expression of these genes was very low in MND samples and very high in the blood of controls, showing a highly significant logFC for these genes. Similarly, *IGFBP5* almost lacked expression in the MND group and had very high expression in the blood of
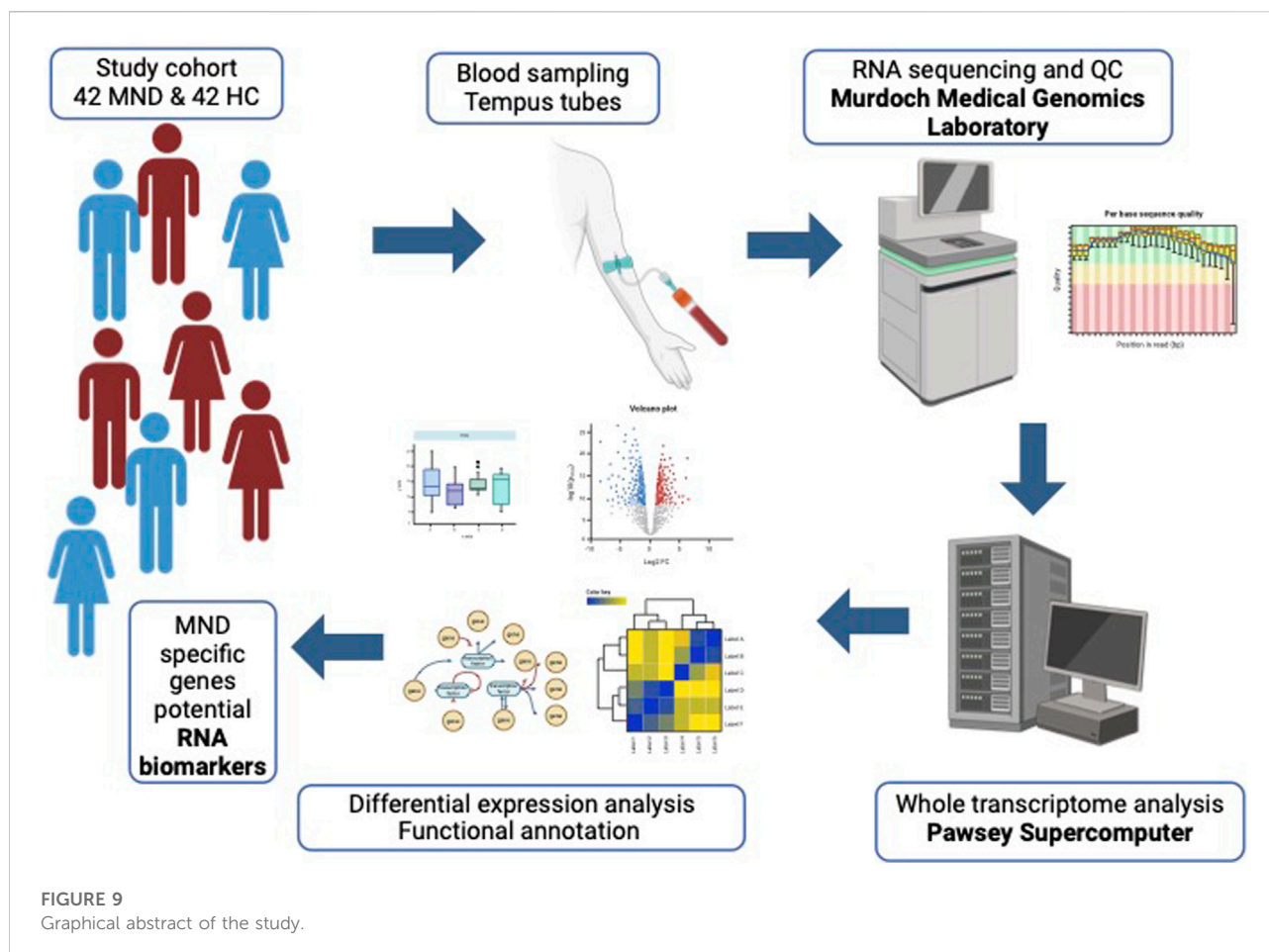
FIGURE 9
Graphical abstract of the study.

control subjects. Overexpression of the *IGFBP5* in mice has induced axonopathy and sensory deficits similar to those seen in diabetic neuropathy [47]. The motor axon degeneration in these mice resembled the pathology seen in MND [47]. *IGFBP5* has been shown to promote neuronal apoptosis in the experimental models and also in patients with spinal muscular atrophy and ALS [48–50].

When discussing these results, we have to consider the effect of MND itself on gene expression and not only the effect of genes on the disease. Most likely, the genes that are significantly downregulated and have very low expression levels in MND patients are the genes that are affected by the MND condition. The cluster of ECM organisation genes indicates the degeneration of the neurones and are the genes directly impacted by the MND. Stanniocalcin 2 (*STC2*) and thrombospondin 2 (*THBS2*) are genes that are related to organogenesis and tissue differentiation [51–53]. Interestingly, the proposed function of these genes is related to collagen genes and MMPs. Therefore, it seems that MND affects tissue reorganisation, and the genes that are required for tissue plasticity are downregulated. We can speculate that genes are not causative for the disease but are affected by the chronic disease condition and lead to enhanced degeneration of neurones.

Conclusion

We performed whole transcriptome analysis from the whole blood RNA and identified 12,972 genes differentially expressed between MND patients and controls. These gene expression changes have the potential to be used as biomarkers to diagnose MND and possibly to evaluate the progression of the disease and drug responsiveness in clinical trials. RNA-based biomarkers have excellent potential as they are quickly responding biomarkers and can be analysed by standardised methods. In conclusion, we were able to identify the characteristic blood gene expression profile of MND patients.

Author contributions

Conceptualization, SK and PT; methodology, SK, KR, MM, and JP; software, SK; validation, SK and AP; formal analysis, SK; investigation, KR, MM, and JP; resources, SK and PT; data curation, SK; writing—original draft preparation, SK; writing—review and editing, SK, KR, MM, JP, AP, and PT;

visualization, SK; supervision, SK; project administration, SK; funding acquisition, SK and PT. All authors contributed to the article and approved the submitted version.

Data availability

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/geo/, GSE277709.

Ethics statement

The studies involving humans were approved by Institutional Review Board of the University of Tartu. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Funding

Acknowledgments

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.ebm-journal.org/articles/10.3389/ebm.2024.10401/full#supplementary-material

SUPPLEMENTARY FIGURE S1
HeatmapMNDEST12972genesBloodRNAseq.

SUPPLEMENTARY FIGURE S2
AllBoxplots.

SUPPLEMENTARY TABLE S1
DescriptionofTheCohort.

SUPPLEMENTARY TABLE S2
Differentially expressed genes.

SUPPLEMENTARY TABLE S3
CompleteListofMNDGENES.

SUPPLEMENTARY TABLE S4
FC_MNDEST_KEGG.

SUPPLEMENTARY TABLE S5
FC_MNDEST_Reactome.

SUPPLEMENTARY TABLE S6
FC_MNDEST_GSEA.

References

1. Siddique T, Deng HX, Ajroud-Driss S. Chapter 132 - motor neuron disease. In: Rimoin D, Pyeritz R, Korf B, editors. *Emery and rimoin's principles and practice of medical genetics*. 6th ed. Oxford: Academic Press (2013). p. 1–22.

2. Brown RH, Al-Chalabi A. Amyotrophic lateral sclerosis. *N Engl J Med* (2017) **377**:162–72. doi:10.1056/nejmra1603471

3. Statland JM, Barohn RJ, McVey AL, Katz JS, Dimachkie MM. Patterns of weakness, classification of motor neuron disease, and clinical diagnosis of sporadic amyotrophic lateral sclerosis. *Neurol Clin* (2015) **33**:735–48. doi:10.1016/j.ncl.2015.07.006

4. Shatunov A, Al-Chalabi A. The genetic architecture of ALS. *Neurobiol Dis* (2021) **147**:105156. doi:10.1016/j.nbd.2020.105156

5. Frohlich A, Pfaff AL, Bubb VJ, Quinn JP, Koks S. Transcriptomic profiling of cerebrospinal fluid identifies ALS pathway enrichment and RNA biomarkers in MND individuals. *Exp Biol Med (Maywood)* (2023) **248**:2325–31. doi:10.1177/15353702231209427

6. Marriott H, Kabiljo R, Hunt GP, Khleifat AA, Jones A, Troakes C, et al. Unsupervised machine learning identifies distinct ALS molecular subtypes in post-mortem motor cortex and blood expression data. *Acta Neuropathol Commun* (2023) **11**:208. doi:10.1186/s40478-023-01686-8

7. Byrne S, Bede P, Elamin M, Kenna K, Lynch C, McLaughlin R, et al. Proposed criteria for familial amyotrophic lateral sclerosis. *Amyotroph Lateral Scler* (2011) **12**:157–9. doi:10.3109/17482968.2010.545420

8. Veldink JH. ALS genetic epidemiology "How simplex is the genetic epidemiology of ALS?" *J Neurol Neurosurg Psychiatry* (2017) **88**:537. doi:10.1136/jnnp-2016-315469

9. Pfaff AL, Bubb VJ, Quinn JP, Koks S. A genome-wide screen for the exonisation of reference SINE-VNTR-alus and their expression in CNS tissues of individuals with amyotrophic lateral sclerosis. *Int J Mol Sci* (2023) **24**:11548. doi:10.3390/ijms241411548

10. Koks G, Pfaff AL, Bubb VJ, Quinn JP, Koks S. At the dawn of the transcriptomic medicine. *Exp Biol Med (Maywood)* (2021) **246**:286–92. doi:10.1177/1535370220954788

11. Pfaff AL, Singleton LM, Kõks S. Mechanisms of disease-associated SINE-VNTR-Alus. *Exp Biol Med (Maywood)* (2022) **247**:756–64. doi:10.1177/15353702221082612

12. Lill M, Koks S, Soomets U, Schalkwyk LC, Fernandes C, Lutsar I, et al. Peripheral blood RNA gene expression profiling in patients with bacterial meningitis. *Front Neurosci* (2013) **7**:33. doi:10.3389/fnins.2013.00033

13. Koks G, Uudelepp ML, Limbach M, Peterson P, Reimann E, Koks S. Smoking-induced expression of the GPR15 gene indicates its potential role in chronic inflammatory pathologies. *The Am J Pathol* (2015) **185**:2898–906. doi:10.1016/j.ajpath.2015.07.006

14. Billingsley KJ, Lättekivi F, Planken A, Reimann E, Kurvits L, Kadastik-Eerme L, et al. Analysis of repetitive element expression in the blood and skin of patients with Parkinson's disease identifies differential expression of satellite elements. *Sci Rep* (2019) **9**:4369. doi:10.1038/s41598-019-40869-z

15. Planken A, Kurvits L, Reimann E, Kadastik-Eerme L, Kingo K, Koks S, et al. Looking beyond the brain to improve the pathogenic understanding of Parkinson's disease: implications of whole transcriptome profiling of Patients' skin. *BMC Neurol* (2017) **17**:6. doi:10.1186/s12883-016-0784-z

16. Zucca S, Gagliardi S, Pandini C, Diamanti L, Bordoni M, Sproviero D, et al. RNA-Seq profiling in peripheral blood mononuclear cells of amyotrophic lateral sclerosis patients and controls. *Sci Data* (2019) **6**:190006. doi:10.1038/sdata.2019.6

17. van Rheenen W, Diekstra FP, Harschnitz O, Westeneng HJ, van Eijk KR, Saris CGJ, et al. Whole blood transcriptome analysis in amyotrophic lateral sclerosis: a biomarker study. *PLoS One* (2018) **13**:e0198874. doi:10.1371/journal.pone.0198874

18. He D, Yang CX, Sahin B, Singh A, Shannon CP, Oliveria JP, et al. Whole blood vs PBMC: compartmental differences in gene expression profiling exemplified in asthma. *Allergy Asthma Clin Immunol* (2019) **15**:67. doi:10.1186/s13223-019-0382-x

19. Moris P, Bellanger A, Ofori-Anyinam O, Jongert E, Yarzabal Rodriguez JP, Janssens M. Whole blood can be used as an alternative to isolated peripheral blood mononuclear cells to measure *in vitro* specific T-cell responses in human samples. *J Immunological Methods* (2021) **492**:112940. doi:10.1016/j.jim.2020.112940

20. Gautam A, Donohue D, Hoke A, Miller SA, Srinivasan S, Sowe B, et al. Investigating gene expression profiles of whole blood and peripheral blood mononuclear cells using multiple collection and processing methods. *PLoS One* (2019) **14**:e0225137. doi:10.1371/journal.pone.0225137

21. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* (2017) **14**:417–9. doi:10.1038/nmeth.4197

22. Love MI, Soneson C, Hickey PF, Johnson LK, Pierce NT, Shepherd L, et al. Tximeta: reference sequence checksums for provenance identification in RNA-seq. *Plos Comput Biol* (2020) **16**:e1007664. doi:10.1371/journal.pcbi.1007664

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

24. Yu G, He QY. ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. *Mol Biosyst* (2016) **12**:477–9. doi:10.1039/c5mb00663e

25. Yu G, Wang LG, Yan GR, He QY. DOSE: an R/Bioconductor package for disease ontology semantic and enrichment analysis. *Bioinformatics* (2015) **31**:608–9. doi:10.1093/bioinformatics/btu684

26. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS: A J Integr Biol* (2012) **16**:284–7. doi:10.1089/omi.2011.0118

27. Kassambra A. Ggpubr: "ggplot2" based publication ready plots (2023).

28. Garau J, Garofalo M, Dragoni F, Scarian E, Di Gerlando R, Diamanti L, et al. RNA expression profiling in lymphoblastoid cell lines from mutated and non-mutated amyotrophic lateral sclerosis patients. *The J Gene Med* (2024) **26**:e3711. doi:10.1002/jgm.3711

29. Mielke JK, Klingeborn M, Schultz EP, Markham EL, Reese ED, Alam P, et al. Seeding activity of human superoxide dismutase 1 aggregates in familial and sporadic amyotrophic lateral sclerosis postmortem neural tissues by real-time quaking-induced conversion. *Acta Neuropathol* (2024) **147**:100. doi:10.1007/s00401-024-02752-8

30. Pokrishevsky E, DuVal MG, McAlary L, Louadi S, Pozzi S, Roman A, et al. Tryptophan residues in TDP-43 and SOD1 modulate the cross-seeding and toxicity of SOD1. *J Biol Chem* (2024) **300**:107207. doi:10.1016/j.jbc.2024.107207

31. Monteiro Neto JR, Ribeiro GD, Magalhães RSS, Follmer C, Outeiro TF, Eleutherio ECA. Glycation modulates superoxide dismutase 1 aggregation and toxicity in models of sporadic amyotrophic lateral sclerosis. *Biochim Biophys Acta (BBA) - Mol Basis Dis* (2023) **1869**:166835. doi:10.1016/j.bbadis.2023.166835

32. Zhao S, Chen R, Gao Y, Lu Y, Bai X, Zhang J. Fundamental roles of the optineurin gene in the molecular pathology of amyotrophic lateral sclerosis. *Front Neurosci* (2023) **17**:1319706. doi:10.3389/fnins.2023.1319706

33. Garofalo M, Pandini C, Bordoni M, Jacchetti E, Diamanti L, Carelli S, et al. RNA molecular signature profiling in PBMCs of sporadic ALS patients: HSP70 overexpression is associated with nuclear SOD1. *Cells* (2022) **11**:293. doi:10.3390/cells11020293

34. Garofalo M, Pandini C, Bordoni M, Pansarasa O, Rey F, Costa A, et al. Alzheimer's, Parkinson's disease and amyotrophic lateral sclerosis gene expression patterns divergence reveals different grade of RNA metabolism involvement. *Int J Mol Sci* (2020) **21**:9500. doi:10.3390/ijms21249500

35. Zhou T, Han Y, Dang Y, Wang X, Zheng YH. A novel HIV-1 restriction factor that is biologically distinct from APOBEC3 cytidine deaminases in a human T cell line CEM.NKR. *Retrovirology* (2009) **6**:31. doi:10.1186/1742-4690-6-31

36. Dang Y, Wang X, Esselman WJ, Zheng YH. Identification of APOBEC3DE as another antiretroviral factor from the human APOBEC family. *J Virol* (2006) **80**:10522–33. doi:10.1128/jvi.01123-06

37. Harris RS, Liddament MT. Retroviral restriction by APOBEC proteins. *Nat Rev Immunol* (2004) **4**:868–77. doi:10.1038/nri1489

38. Liang W, Xu J, Yuan W, Song X, Zhang J, Wei W, et al. APOBEC3DE inhibits LINE-1 retrotransposition by interacting with ORF1p and influencing LINE reverse transcriptase activity. *PLoS One* (2016) **11**:e0157220. doi:10.1371/journal.pone.0157220

39. Pfaff AL, Bubb VJ, Quinn JP, Koks S. Locus specific reduction of L1 expression in the cortices of individuals with amyotrophic lateral sclerosis. *Mol Brain* (2022) **15**:25. doi:10.1186/s13041-022-00914-x

40. Takei H, Fukuda H, Pan G, Yamazaki H, Matsumoto T, Kazuma Y, et al. Alternative splicing of APOBEC3D generates functional diversity and its role as a DNA mutator. *Int J Hematol* (2020) **112**:395–408. doi:10.1007/s12185-020-02904-y

41. Stenglein MD, Burns MB, Li M, Lengyel J, Harris RS. APOBEC3 proteins mediate the clearance of foreign DNA from human cells. *Nat Struct Mol Biol* (2010) **17**:222–9. doi:10.1038/nsmb.1744

42. Stenglein MD, Harris RS. APOBEC3B and APOBEC3F inhibit L1 retrotransposition by a DNA deamination-independent mechanism. *J Biol Chem* (2006) **281**:16837–41. doi:10.1074/jbc.m602367200

43. Liu EY, Russ J, Cali CP, Phan JM, Amlie-Wolf A, Lee EB. Loss of nuclear TDP-43 is associated with decondensation of LINE retrotransposons. *Cel Rep* (2019) **27**:1409–21.e6. doi:10.1016/j.celrep.2019.04.003

44. Irwin KE, Jasin P, Braunstein KE, Sinha IR, Garret MA, Bowden KD, et al. A fluid biomarker reveals loss of TDP-43 splicing repression in presymptomatic ALS-FTD. *Nat Med* (2024) **30**:382–93. doi:10.1038/s41591-023-02788-5

45. La Cognata V, Gentile G, Aronica E, Cavallaro S. Splicing players are differently expressed in sporadic amyotrophic lateral sclerosis molecular clusters and brain regions. *Cells* (2020) **9**:159. doi:10.3390/cells9010159

46. Kletzl H, Marquet A, Günther A, Tang W, Heuberger J, Groeneveld GJ, et al. The oral splicing modifier RG7800 increases full length survival of motor neuron 2 mRNA and survival of motor neuron protein: results from trials in healthy adults and patients with spinal muscular atrophy. *Neuromuscul Disord* (2019) **29**:21–9. doi:10.1016/j.nmd.2018.10.001

47. Rauskolb S, Dombert B, Sendtner M. Insulin-like growth factor 1 in diabetic neuropathy and amyotrophic lateral sclerosis. *Neurobiol Dis* (2017) **97**:103–13. doi:10.1016/j.nbd.2016.04.007

48. Wilczak N, de Vos RA, De Keyser J. Free insulin-like growth factor (IGF)-I and IGF binding proteins 2, 5, and 6 in spinal motor neurons in amyotrophic lateral sclerosis. *Lancet* (2003) **361**:1007–11. doi:10.1016/s0140-6736(03)12828-0

49. Kaymaz AY, Bal SK, Bora G, Talim B, Ozon A, Alikasifoglu A, et al. Alterations in insulin-like growth factor system in spinal muscular atrophy. *Muscle Nerve* (2022) **66**:631–8. doi:10.1002/mus.27715

50. Guo S, Lei Q, Yang Q, Chen R. IGFBP5 promotes neuronal apoptosis in a 6-OHDA-toxicant model of Parkinson's disease by inhibiting the sonic hedgehog signaling pathway. *Med Princ Pract* (2024) **33**:269–80. doi:10.1159/000538467

51. Qu HL, Hasen GW, Hou YY, Zhang CX. THBS2 promotes cell migration and invasion in colorectal cancer via modulating Wnt/β-catenin signaling pathway. *The Kaohsiung J Med Sci* (2022) **38**:469–78. doi:10.1002/kjm2.12528

52. Lai R, Ji L, Zhang X, Xu Y, Zhong Y, Chen L, et al. Stanniocalcin2 inhibits the epithelial-mesenchymal transition and invasion of trophoblasts via activation of autophagy under high-glucose conditions. *Mol Cell Endocrinol* (2022) **547**:111598. doi:10.1016/j.mce.2022.111598

53. Zhou J, Li Y, Yang L, Wu Y, Zhou Y, Cui Y, et al. Stanniocalcin 2 improved osteoblast differentiation via phosphorylation of ERK. *Mol Med Rep* (2016) **14**:5653–9. doi:10.3892/mmr.2016.5951