scientific reports

OPEN

Transcriptomic analysis reinforces the implication of spatacsin in neuroinflammation and neurodevelopment

Liriopé Toupenet Marchesi[1,2], Daniel Stockholm[2,3], Typhaine Esteves[1,2], Marion Leblanc[1,2], Nicolas Auger[1,2], Julien Branchu[1], Khalid Hamid El Hachimi[1,2] & Giovanni Stevanin[2,4]✉

Hereditary spastic paraplegia (HSP) encompasses a group of rare genetic diseases primarily affecting motor neurons. Among these, spastic paraplegia type 11 (SPG11) represents a complex form of HSP caused by deleterious variants in the *SPG11* gene, which encodes the spatacsin protein. Previous studies have described several potential roles for spatacsin, including its involvement in lysosome and autophagy mechanisms, neuronal and neurites development or mitochondria function. Despite these findings, the precise function of the spatacsin protein remains elusive. To elucidate its function, we conducted an extensive RNA sequencing (RNAseq) experiment and transcriptomic analysis in three distinct neural structures (cerebellum, cortex and hippocampus) and at three different ages (6 weeks, 4 months and 8 months) in both wild type and *Spg11⁻/⁻* mice. Our functional analysis of differentially expressed genes (DEGs) and Gene Set Enrichment Analysis (GSEA) revealed dysregulation in pathways related to inflammation, RNA metabolism and neuronal and neurite development, factors frequently implicated in neurodegenerative disorders. Notably, we also observed early deregulation in cellular pathways related to cell proliferation. Our results represent a significant step towards a better understanding of the functions of spatacsin in the cell and the underlying cellular mechanisms disrupted by its absence.

Keywords Spastic paraplegia, Transcriptomic analysis, Mouse model, SPG11, DEG, GSEA

Abbreviations
ALR   Autophagic lysosome reformation
ALS   Amyotrophic lateral sclerosis
Chr2   Chromosome 2
CNS   Central nervous system
DA   Differential analysis
DEG   Differentially expressed gene
FC   Fold change
FDR   False discovery rate
GO   Gene ontology
GOBP   Gene ontology biological process
GOCC   Gene ontology cellular component
GOMF   Gene ontology molecular function
GSEA   Gene set enrichment analysis
HSP   Hereditary spastic paraplegia
KEGG   Kyoto encyclopedia of genes and genomes
MsigDB   Molecular signature database
PCA   Principal component analysis
SPG11   Spastic paraplegia type 11
t-SNE   t-Distributed stochastic neighbor embedding algorithm

[1]Paris Brain Institute (ICM), Sorbonne University, INSERM, CNRS, APHP, Paris, France. [2]PSL Research University, EPHE, Paris, France. [3]Sorbonne Université, INSERM, Centre de Recherche Saint-Antoine, CRSA, Paris, France. [4]Institut des Neurosciences cognitives et intégratives d'Aquitaine (INCIA), Bordeaux University, CNRS, Bordeaux, France. ✉email: giovanni-b.stevanin@inserm.fr

Hereditary spastic paraplegias (HSP) constitute a large family of neurological disorders affecting particularly the motor neurons. They are characterized by a lower-limb spasticity and weakness due to degeneration in the corticospinal tracts[1–3]. The spastic paraplegia type 11 (SPG11) is due to pathogenic mutation in the *SPG11* gene and it is one of the most frequent forms of autosomal recessive HSP[3–5]. In addition to the spastic gait disorder that is common to this family of neurological disorders, patients present frequently other neurological signs such as cognitive impairment, peripheral neuropathy, retinopathy and cerebellar ataxia, often associated to a thinning of the corpus callosum at brain MRI[6–9].

To study the underlying pathological mechanisms due to the absence of the full-length protein product of the *SPG11* gene, spatacsin, an *Spg11* knockout mouse model was previously generated and characterized[10]. The *Spg11*−/− mouse displays early-onset motor and cognitive deficits mimicking the human disease. Indeed, the *Spg11*−/− mouse has a motor deficit from 6 weeks of age and first memory alterations are observed at 4 months. Moreover, atrophy of hippocampus has been observed from 4 months, followed at 8 months by atrophy of the cortex and cerebellum[10]. *SPG11* encodes for spatacsin, a large protein highly conserved across vertebrates[6]. Despite the fact that the exact function of spatacsin is not well understood, several studies indicate that this protein could be involved in autophagy-lysosomal machinery[10–17], mitochondrial dysfunctions[18,19] and neurodevelopment[18,20–24].

So far, no transcriptomic analysis has been performed on a SPG11 mouse model. Transcriptomic analysis allows a more holistic approach compared to studies targeting particular cellular mechanisms. Indeed, transcriptomic data offers an overview of gene expression, and therefore, of tissue functioning at a given time. The identification of cellular pathways disturbed in the *Spg11*−/− mouse could not only strengthen current hypotheses on the functions of spatacsin, but also provide new insights into the cellular processes disturbed by its absence. In the present study, we performed a large RNA-seq experiment on 90 samples from 15 mice of each genotype (*Spg11*+/+ and *Spg11*−/−) covering three different ages (6 weeks, 4 months and 8 months) and three different nerve structures (cerebellum, hippocampus and cortex). In the *Spg11*−/− mouse model, atrophy of these three brain structures has been described. In addition, the ages studied cover the major stages of the disease, from the first symptoms appearance to generalized brain atrophy[10]. Thus, the scope of the study is expected to provide us information on the gene deregulations dynamics at work, at different stages of the disease and in the main brain structures affected.

Our results confirm the involvement of pathologic mechanisms identified in previous studies in *Spg11*−/− mice. Notably, we observed deregulations in inflammation pathways particularly at 8 months. In addition, analysis of transcriptomic data highlighted the early character of neurodegeneration in *Spg11*−/− mice. Indeed, cellular pathways frequently deregulated in neurodegenerative diseases, such as those involved in RNA metabolic process or neuronal and synaptic development, are disrupted from 4 months of age in *Spg11*−/− mice. Interestingly, we also report an unexpected early deregulation in pathways related to cellular proliferation.

Materials and methods

Mouse

The *Spg11*-knockout (KO) mouse was previously described[10]. Briefly, the gene-targeting vector was constructed by inserting two successive STOP codons into the exon 32 of the *Spg11* gene, mimicking mutations found in most patients. The vector was transfected in MCI-129 Sv/Pas ES cells and homologous recombinant clones were injected into blastocysts. Then chimeric mice were mated with C57BL/6 animals. Finally, heterozygous mice were backcrossed with C57BL/6 N mice for 11 generations to stabilize the genome in the novel background before this study. The mice studied are all males.

Tissue collection and RNA extraction

We selected three representative regions of the Central Nervous System (CNS) affected during the course of the disease (cerebellum, cortex and hippocampus), as described previously[10], and sampled them at three different ages: 6 weeks which represented the beginning of the motor symptoms, 4 months being a fully symptomatic stage and 8 months when neurodegeneration occurs.

The tissues from five biological replicates at each condition were collected and frozen in liquid nitrogen. Tissue was homogenized using a FastPrep bead homogenizer (MP Biomedicals, LLC) in tubes containing Lysing Matrix D beads (MP Biomedicals 6913–100) and lysis buffer (Maxwell SimplyRNA Tissue - Promega kit). Subsequently, RNA extraction was performed in a Maxwell RSC instrument (AS4500) by transferring the tissue lysates to the first well of each RSC cartridge pre-filled with buffer (Maxwell SimplyRNA Tissue kit). Total RNAs were eluted in 60 to 70 µl of nuclease-free water. RNA quality was measured using the Agilent TapeStation system (minimal requirement: RNA Integrity Number > 8).

Libraries preparation, sequencing and alignment

Libraries for RNA-seq were prepared from 1 µg of each RNA extract using the KAPA mRNA hyperprep kit (Illumina). The quality of the libraries was checked by migration on the Labchip GX capillary electrophoresis system (Perkin Elmer). The quantity of cDNA was measured using the QuBit dsDNA BR Assay kit (ThermoScientific). Sequencing was performed using an Illumina NextSeq500 sequencer (Illumina, San Diego, CA) providing from 40,839,430 to 102,986,634 reads per sample. Base call files were converted to the FASTQ format using bcl2fastq software. FastQC software was used to check the quality of raw data. All reads were then aligned using STAR software (2.4.0f1) to the mouse reference genome (starindex_mm10). The count matrix was generated from BAM files with the package featureCounts (v1.6.2) from Rsubread[25].

Data analysis

Following analysis was conducted with the language R (v4.2.2) (R Core Team 2023). The package EdgeR was used for the analysis of the differential expression between mouse samples from the count matrix[26]. Genes with low expression were removed on a basis of a total count of less than 15 and a minimum count of 10 per sample. Gene counts were normalized by applying the trimmed mean of M-values (TMM) normalization method[27]. We conducted nine independent differential analysis (comparison) by comparing five $Spg11^{-/-}$ mice vs. five $Spg11^{+/+}$ mice for each type of structure (cerebellum, hippocampus and cortex) at each age (6 weeks, 4 months and 8 months).

Differential analyses were performed with general linear model framework likelihood ratio test. Multiple hypothesis-adjusted p-values were calculated with the Benjamini-Hochberg procedure to control FDR (False Discovery Rate). For each gene, 18 parameters (9 FDR and 9 Fold Change values) were obtained. Differentially expressed genes (DEGs) for each of the nine comparisons were defined as genes with an absolute value of log2 fold change (logFC) greater than 0.5 and a FDR less than 0.05. Volcano plots, heatmaps, Venn diagrams, and genomic charts were generated with the following packages: ggplot2[28], gplots, ggVennDiagram[29], GenomicRanges[30], and Gviz[31].

For the identification of enriched pathways, we conducted two analyses: one analysis by over-representation and one Gene Set Enrichment Analysis (GSEA). For the first one, the choice was made to select genes with at least one FDR $< 10^{-4}$ among the nine comparisons studied. We obtained a total of 999 genes that were visualized by t-SNE (t-Distributed Stochastic Neighbor Embedding algorithm)[32] using the Rtsne package. The threshold was chosen to balance statistical stringency with practical visualization needs, ensuring a manageable number of data points for 2D representation using t-SNE dimensionality reduction.

GSEA was conducted with the clusterProfiler package (v3.14.3)[33] on KEGG (Kyoto Encyclopedia of Genes and Genomes) and MsigDB (Molecular Signature database) Hallmark databases. The p-value was confirmed with the Benjamini–Hochberg correction procedure. Only the pathways with p-adjusted lower than 0.05 were firstly selected. If a pathway had a p-adjusted < 0.05 in at least one condition, it was considered in other condition with a p-adjusted until 0.25 according to the recommendations of the GSEA developers[34].

Results

RNAseq analysis

We performed a large-scale bulk RNA sequencing (RNA-seq) experiment on the $Spg11^{-/-}$ mouse model. Robust data were obtained on five biological replicates of each condition that included two genotypes ($Spg11^{+/+}$, $Spg11^{-/-}$) covering three critical ages during the pathological process (6 weeks: first motor troubles, 4 months: memory and motor dysfunctions without neuronal loss, 8 months: neurodegeneration) on three CNS structures (cerebellum, cortex and hippocampus). The number of raw reads obtained averaged 70 million per sample (range 40,839,430 to 102,986,634) out of which 93% could be confidently mapped to the *mus musculus* genome (starindex_mm10). This primary step enabled the assembly of 55,573 non-redundant transcripts and their quantification across all samples. After filtration and normalization steps, we obtained a set of 25,421 transcripts expressed in at least one condition.

We first performed a principal component analysis (PCA) order to get insights in samples homogeneity. The PCA was visualized on two dimensions with principal component 1 and 2 as axes (Figure S1). The two first components appeared to be strongly correlated to the type of brain structure. Indeed, each nerve structure forms a cluster, although the hippocampus and motor cortex are closer (Figure S1A). This result suggests that the three nerve structures have distinct gene expression profiles, as expected, without mis-sampling. Furthermore, by plotting the PCAs for each condition, we can note that the two genotypes are clearly segregated at 4 months and 8 months in the three brain structures. This separation is weaker at 6 weeks showing a milder effect of the mutation at this stage, as expected (Figure S1B).

Dynamics of gene expression dysregulation

We performed a differential analysis (DA) between $Spg11^{-/-}$ and $Spg11^{+/+}$ mice for each of the nine conditions: cerebellum, cortex and hippocampus at 6 weeks, 4 months and 8 months. Data were normalized and filtered with the EdgeR package[26]. We established cut off values for the two main parameters: False Discovery Rate (FDR) and Fold Change (FC). The conditions to filter out differentially expressed genes (DEGs) were: FDR < 0.05 and |log(FC)|>0.5.

Figure 1 shows the volcano plot of the DEGs in the nine conditions. We can note that the number of DEGs is clearly lower at 6 weeks than at the other ages studied in every structure. At 4 months, the number of DEGs is the most pronounced for the three structures. Moreover, in accordance with PCA results, the expression profile of the DEGs from the cerebellum diverges from those of the hippocampus and the cortex. Indeed, while the number of DEGs decreases at 8 months in the latter, it is stable in the cerebellum with a strong increase in up-regulated genes (Fig. 1).

To complement the overview provided by the volcano plot, Table 1 describes the DA results. In the cerebellum, the number of DEGs was 33 (18 up-regulated and 15 down-regulated), 283 (76 up-regulated and 207 down-regulated) and 266 (190 up-regulated and 76 down-regulated) at 6 weeks, 4 months and 8 months, respectively. In the cortex, this count moved from 20 genes differentially expressed (11 up-regulated and 9 down-regulated), to 635 genes (210 up-regulated and 425 down-regulated) and 115 genes (46 up-regulated and 69 down-regulated) at the three same ages. Finally, in the hippocampus, at 6 weeks there was 38 DEGs (22 up-regulated and 16 down-regulated), then 321 DEGs at 4 months (179 up-regulated and 142 down-regulated) and 131 DEGs at 8 months (68 up-regulated and 63 down-regulated).

The table presents the number of genes included in the differential analysis after the filtration steps and the number of DEGs in the cerebellum, cortex and hippocampus of $Spg11^{-/-}$ mice compared to $Spg11^{+/+}$ mice at 6
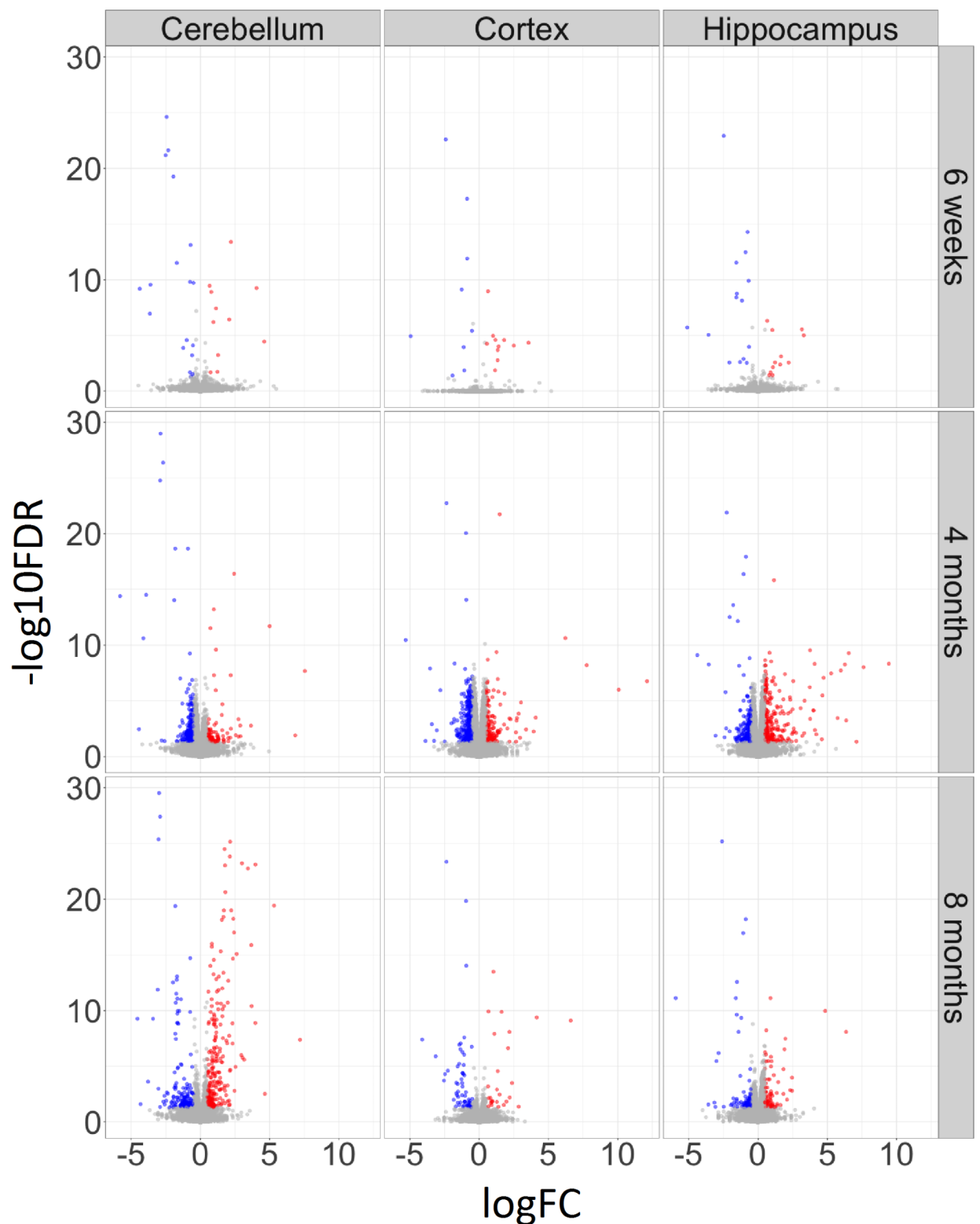
Fig. 1. Volcano plot illustrating differential gene expression analysis between *Spg11*$^{-/-}$ and *Spg11*$^{+/+}$ mice under nine conditions. The x-axis represents log2 fold change (logFC) while the y-axis shows the negative log10 FDR (-log10FDR). To be considered as differentially expressed a gene must have a |logFC|>0.5 and an FDR<0.05. Up-regulated genes are in red and down-regulated genes in blue.

Structure	Cerebellum			Cortex			Hippocampus		
Age	6w	4m	8m	6w	4m	8m	6w	4m	8m
Nb of genes included in DA	19 470	19 450	19 699	19 426	18 992	19 562	19 442	19 218	19 493
Nb of DEGs ⬆⬇	**33** 18/15	**283** 76/207	**266** 190/76	**20** 11/9	**635** 210/425	**115** 46/69	**38** 22/16	**321** 179/142	**131** 68/63

Table 1. Results of the differential analysis (DA) between *Spg11*[−/−] and *Spg11*[+/+] mice.

weeks, 4 months and 8 months. To be considered as differentially expressed a gene must have a |log(FC)|>0.5 and an FDR<0.05. The number of up regulated genes is written in red, and the number of down regulated genes is written in blue.

Deregulated pathways according to the mouse phenotype

We first validated the RNAseq data focusing on known deregulations in the disease. At the cellular level, one of the major hallmarks of SPG11 disease is autophagy-lysosomal system impairment. Indeed, several lysosomal defects have been described in our model of *Spg11*[−/−] mouse (Branchu et al. 2017) and in a similar model by others[10,15]. Accumulation of lysosomal or autophagosome structures have been observed in various cell lines including fibroblasts from SPG11 patients[14,16] and from *Spg11*[−/−] mouse model both in vivo in brain sections or in vitro in Mouse Embryonic Fibroblasts[10,15,17]. Moreover, spatacsin is known to be involved in Autophagic Lysosome Reformation (ALR)[14]. We therefore expected a deregulation of the two GO term related to autophagy and lysosome, GO: 0006914 and GO: 0005764 respectively.

Figure 2 shows the deregulated genes related to autophagy and lysosome pathways in the nine conditions. As expected, the expression of genes in these pathways is modified in *Spg11*[−/−] mice, following a pattern similar to that observed on all DEGs, with minimal deregulation at six weeks and a peak at 4 months. However, this dynamic contrasts with histological findings in *Spg11*[−/−] mice[10], which show a progressive accumulation of lysosomal material throughout the animal's life, suggesting worsening lysosomal and autophagic alterations with age, a progression only partially reflected in our transcriptomic data.

Although the analyses did not reveal any significantly deregulated sub-pathways within the autophagy and lysosome pathways (Table S1), certain DEGs are involved in specific sub-pathways. In the lysosomal pathway, changes were observed in the cortex at four months, including genes involved in lysosomal transport (*Grn*, *Hgs*, *Lamp2*, *Psap*, and *Scarb2*) and sphingolipid metabolism (*Abca2*, *Arsa*, *Asah1*, *Ctsa*, *Galc*, *Hexa*, *Hexb*, *Neu1*, *Psap*, and *Smpd1*). Moreover, in the cerebellum at eight months, we observe a large cluster of up-regulated genes and notably genes associated with macrophage activation (*App*, *Cst7*, *Ctsc*, *Grn*, *Ldlr*, and *Slc11a1*). Regarding autophagy-related pathways, we can notice, at four months, pathways involved in TOR signaling (*Clec16a*, *Pink1*, *Prkaa1*, *Prkaa2*, *Stk11*, and *Wdr24*), or in the response to starvation (*Ambra1*, *Bcas3*, *Clec16a*, *Lamp2*, *Map1lc3*, *Prkaa1*, *Prkaa2*, *Ulk1*, *Ulk2*, and *Wdr24*), as well as in the ESCRT machinery (*Chmp2b*, *Chmp4c*, *Chmp5*, *Chmp6*, and *Vps4b*).

Canonical pathways enriched for DEGs

To identify more complicated spatiotemporal relationships among genes, we selected genes with a FDR<10^{-4} in at least one of the nine conditions, i.e. 999 genes, and used t-SNE representation[32] shown in Fig. 3 to reduce the dimensions and visualize the clusters. We identified a total of eight clusters, but number of genes with an FDR<10^{-4} vary greatly between conditions. We conducted an over-representation analysis on the selected genes to identify the significantly deregulated pathways for each of the nine conditions. Table S2 summarizes the deregulated pathways referenced later in the text. Furthermore, the supplementary Table S3 provides a comprehensive list of all deregulated pathways for each of the nine conditions.

Only two clusters are common in the nine conditions studied, one containing up-regulated genes and one with down-regulated genes (clusters A and B), both are enriched with genes localized on *Spg11* locus region of chromosome 2.

Regarding the cerebellum, at 6 weeks, 24 genes have a FDR<10^{-4}, extending to 125 genes at 4 months. Only the two clusters described above (clusters A and B) are clearly segregated in these conditions. However, at 8 months we have 174 genes with an FDR<10^{-4} that are separated in 4 clusters: the two clusters A and B and two additional clusters (clusters C and D). The cluster C is composed of 116 up-regulated genes and highly enriched in genes involved in immune/inflammatory response. The cluster D is constituted of 17 genes including six mitochondrial genes and five involved in oxidative phosphorylation (*mt-Atp6*, *mt-Co3*, *mt-Nd4l*, *mt-Co2* and *mt-Atp8*).

In the cortex, at 6 weeks, only 16 genes have an FDR<10^{-4}, which aggregate in the A and B clusters. At 4 months 559 genes have an FDR<10^{-4} which is the condition with the most DEGs in our t-SNE analysis. In addition to clusters A and B, the genes were condensed in two main clusters (clusters E and F) and an additional
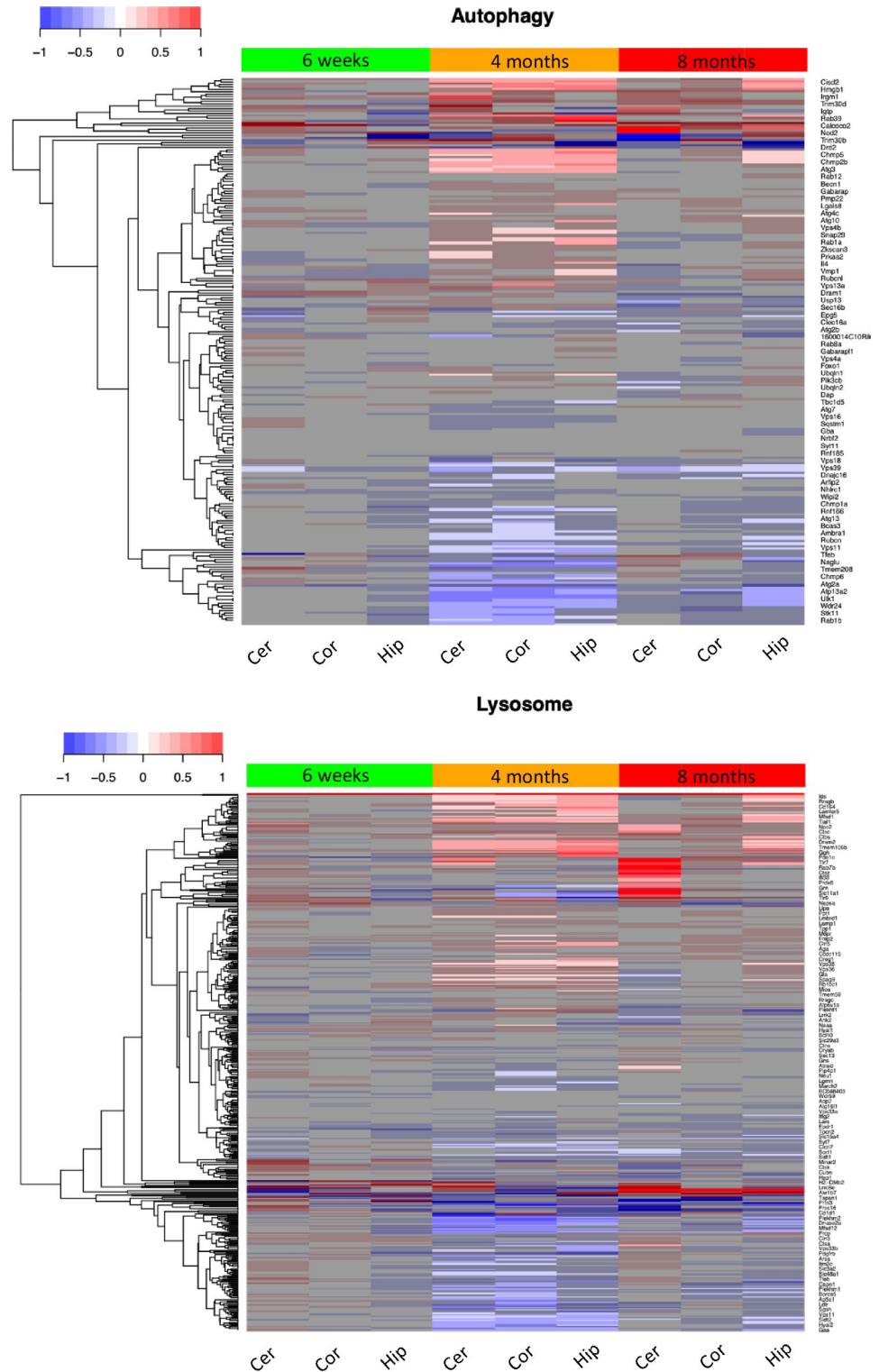
Fig. 2. Heatmaps showing all expressed genes belonging to the autophagy (GO:0006914) and lysosome (GO: 0005764) pathways for the nine conditions studied. Downregulated genes (*Spg11*$^{+/+}$ vs. *Spg11*$^{-/-}$) are in blue and upregulated genes are in red. The shaded boxes correspond to genes that do not subscribe to AD criteria (|logFC|>0.5 and p-value < 0.05). Thus, only unshaded blue or red genes are differentially expressed genes.

cluster localized between them (cluster G). For the latter we could not identify a particular enrichment. The cluster E contains down-regulated genes expressed in several pathways. Among the most enriched pathways we can notice pathways involved in synapse formation, as well as, pathways relative to cellular projection and neurodevelopment. Interestingly, this cluster contains also genes expressed in pathways involved in lipid
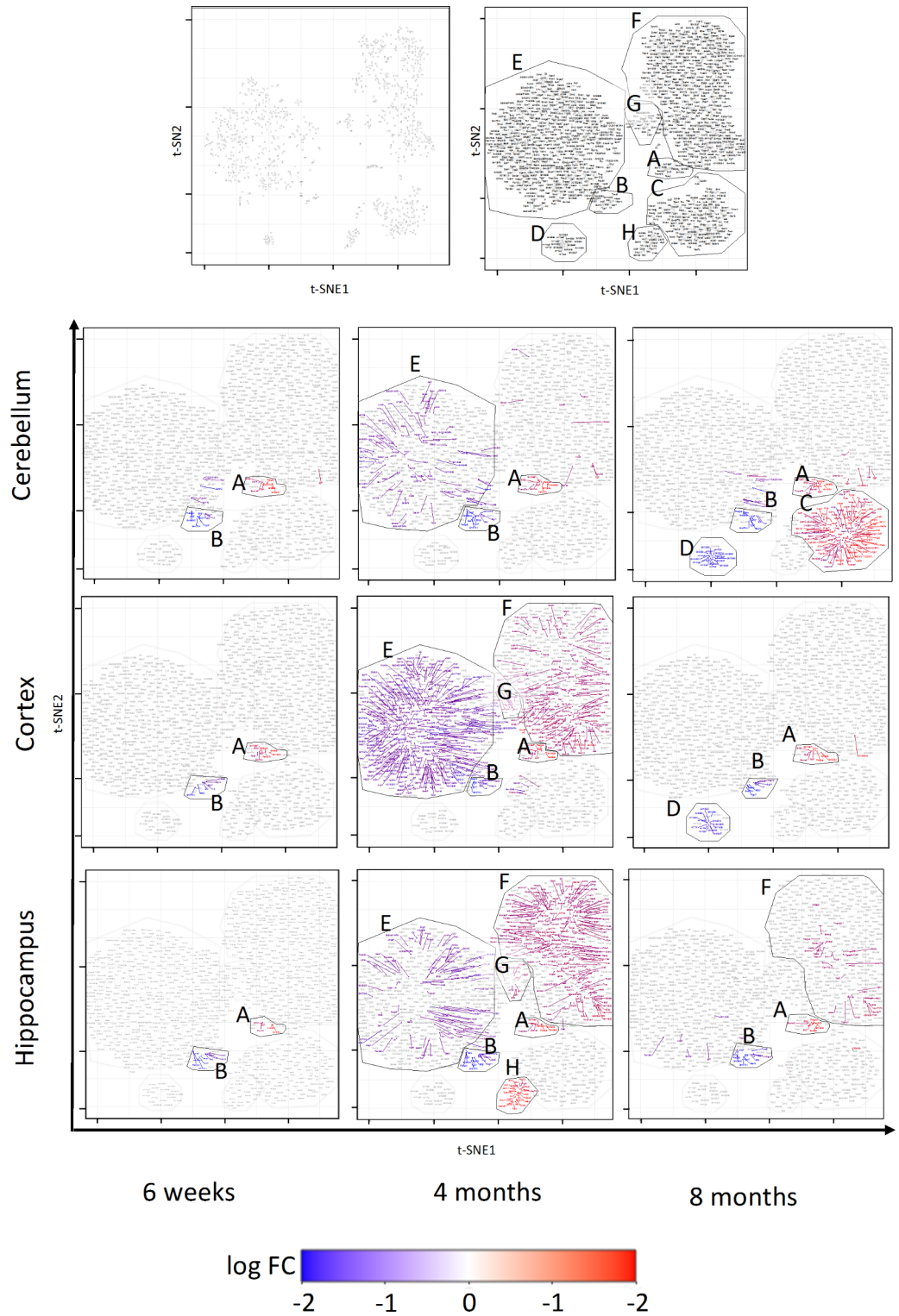
Fig. 3. t-SNE clustering analysis for the nine conditions studied where each point is one of the 999 genes with a FDR < 10^{-4} in at least one of the conditions. The first row shows raw t-SNE, one without names (left) and one with gene names as well as manually delineated clusters and their corresponding labels (right). The following rows emphasize each condition by indicating in false colors how are deregulated gene. Blue indicates genes down regulated, red indicates genes up regulated and grey indicates genes with a FDR > 10^{-4} for the considerate condition. The two clusters A and B contain genes belonging to *Spg11* flanking region. Cluster C is mainly composed of up regulated genes involved in immune response. Cluster D contains mitochondrial genes. Cluster E is made up of genes belonging to pathways relating to neurodevelopment, synapse formation and to a lesser extent lipid metabolism. Cluster F is composed of genes involved in RNA process and intracellular transport. At last, clusters G and H have no predominant pathway enrichment.

metabolism. Concerning the cluster F, it contains 174 up-regulated genes and it is highly enriched in genes engaged in RNA process, including 57 involved in RNA binding. We can also find some genes related to intracellular transport. Surprisingly, at 8 months only 32 genes have an FDR $< 10^{-4}$ and they are grouped into three clusters. The two clusters (A and B) containing genes flanking the *Spg11 locus* and the cluster D containing mitochondrial genes, also deregulated at 8 months in the cerebellum, including *mt-Atp6*, *mt-Co3*, *mt-Co2* and *mt-Atp8* involved in oxidative phosphorylation.

Hippocampus has a similar pattern as cortex. Indeed, very few genes have an FDR $< 10^{-4}$ at 6 weeks and 8 months; respectively 16 and 64. At 6 weeks, the spotted genes exclusively belongs to clusters A and B whereas in 8 months we can notice 36 genes that are members of the cluster F and more specially eight of them are localized in endosomes: *Rap1a*, *Rap2c*, *Hmgb1*, *Tgfbr1*, *Abhd17b*, *Trappc2*, *Atp6ap2*, *Rnf11*. At 4 months, 391 genes have an FDR $< 10^{-4}$ distributed in six clusters, the two clusters A and B, two smaller clusters including one specific at this condition (clusters G and H) and the two large clusters E and F. Concerning the clusters G and H, one is composed of 25 genes and the other of 10 genes, but no pathway enrichment was detected. Compared to cortex at 4 months, less genes (95 genes) are highlighted in the cluster E, but we note similar enriched pathways relative to synapse and neuron projection, but not for lipid metabolism pathways in contrast to the cortex. Regarding the cluster F, 235 genes have an FDR $< 10^{-4}$ at hippocampus 4 months. Similarly to cortex, a fraction of these genes is involved in RNA binding although to a lesser extent (32 genes). On the other hand, genes involved in cellular transport are more numerous (44 genes). In particular, pathways link with endosome formation and transport appears deregulated.

Positional effect of gene dysregulation

For each CNS structure independently, we have identified the DEGs common to the three ages studied. There are only 21 DEGs shared by the three ages for the cerebellum, 15 DEGs for the cortex and 18 for the hippocampus, representing 4%, 2.2% and 4.4% of the total DEGs per structure, respectively. Moreover, for the three structures, the majority of DEGs are only deregulated at a single age. Indeed, between 4 months and 8 months the number of common DEGs remains low: 1.9% for the cerebellum, 6.8% for the cortex and 9% for the hippocampus. We note once again that the cerebellum follows a different pattern from the other two structures with a particularly low number of shared DEGs at the three different ages (Fig. 4). The table S4 details the DEGs shared across the three ages for each structure.
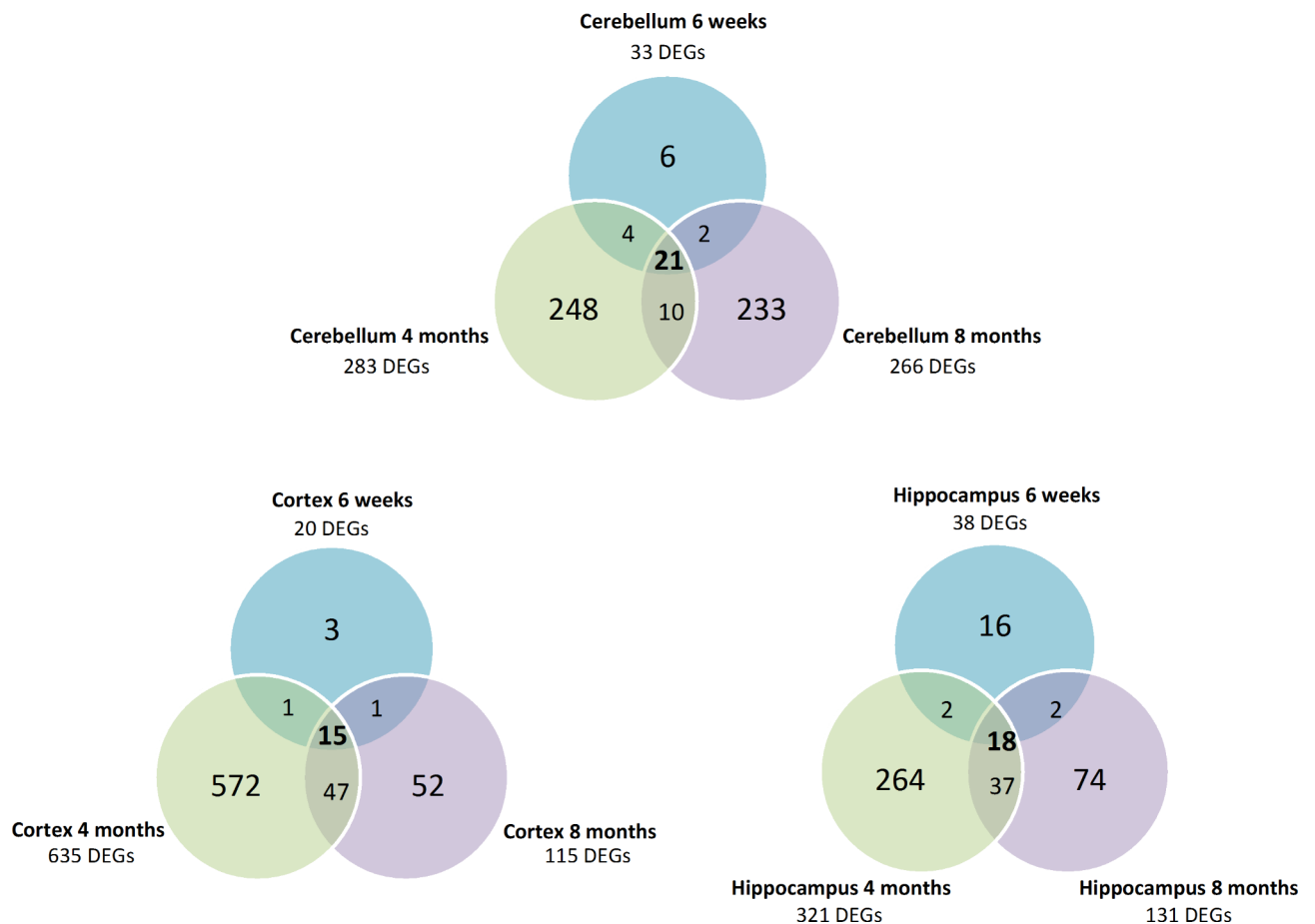

Fig. 4. Venn diagrams showing the commons DEGs (*Spg11*$^{+/+}$ vs. *Spg11*$^{-/-}$) for three ages in the cerebellum, cortex and hippocampus.

Outstandingly, the majority of DEGs (22/27 genes) shared by all conditions are genes located in the region of the *Spg11* gene on chromosome 2 (Table 2). This result strongly suggests that their deregulation is not a direct consequence of the *SPG11* gene expression decrease. Of note, a local rearrangement was excluded by sequencing so that either this phenomenon is the consequence of local gene regulations or due to the way the strain was generated. Indeed, the importance of genetic background in a phenotype exhibited by transgenic mice is well established, that's why a backcrossing strategy is commonly implemented. This procedure will eliminate most donor material on unlinked chromosomes. But the main problem is the region flanking the targeted gene that will "travel" or segregate with the selected locus and therefore remain originating from the donor in a proportion that will depend on the number of backcrosses. This is called gene passenger effect[35,36]. As previously described[10], the *Spg11* gene was inactivated by a gene-targeting vector transfected in MCI-129 Sv/Pas ES cells line. Homologous recombinant clones were injected into blastocysts and then chimeric mice were mated with C57BL/6 animals. Finally, heterozygous mice were backcrossed with C57BL/6 N mice for 11 generations. Thus, the flanking region of the *Spg11* gene in the *Spg11*$^{+/+}$ mice comes from the C57BL/6 strain, while it is predictably originating from the 129-derived ESCs in the *Spg11*$^{-/-}$ mice for an approximately 1,7 M bp after 11 backcrossing and selection of the mutant carriers.

The potential passenger gene effect can be visualized by highlighting the DEGs in the genome (Figure S2). The majority of them are in the region of the *Spg11* locus. We plotted a second volcano plot highlighting the genes belonging to *Spg11* locus (Figure S3). It appears that the DEGs at 6 weeks are mainly genes flanking the *Spg11* locus (cerebellum 21/33; cortex 13/20 and hippocampus 18/38), which is not the case for others conditions (Table S5). Among the DEGs flanking the *Spg11* locus, there are 8 DEGs (*Bloc1s6*, *Bub1b*, *Eif3j1*, *Ell3*, *Exd1*, *Fam227b*, *Ivd*, *Pla2g4e*) found in the nine conditions studied. They are located on chromosome 2 between positions chr2: 118,428,692 and chr2: 125,993,924. We decided to exclude the genes included in this region for the GSEA analysis hypothesizing that most of their deregulation is probably due to the passenger gene effect.

GSEA: new pathways identification

To continue the analysis of the transcriptomic data, we performed a GSEA analysis without taking into account the DEGs flanking the *Spg11* region on chromosome 2 potentially resulting from the passenger effect. In our previous analysis, data processing associated a FC and a FDR to each gene. In over-representation analysis, these two values are used to establish arbitrary cut-offs in order to determine which gene is differentially expressed or not. Conversely, in GSEA analysis, all genes are considered regardless of their deregulation or not, and the FC and FDR are used only to rank the genes. Then, the goal is to determine whether a subset of genes belonging to a pathway tend to occur toward the top (or the bottom) of the ranked genes[37]. The advantage of GSEA analysis is to not focus on individual genes but on biological processes and cellular pathways, and therefore, allows finding the accumulated effects of gene expression in two different phenotypes.

For our analysis, we used the curated gene sets from the KEGG Pathways database[38] and the Hallmark MsigDB database[34]. The results of GSEA analysis are summarized in Tables 3 and 4. We identified 34 MsigDB pathways and 27 KEGG pathways with a p-value adjusted (p-adjusted) < 0.05 in at least one of the nine conditions. None of the tested pathways was significantly deregulated in all conditions. Interestingly, the GSEA results revealed numerous deregulated pathways at the earliest tested stage (6 weeks).

The immune/inflammatory response is clearly altered in our model. In cerebellum, GSEA results show a strong deregulation of immune response pathways from 6 weeks, intensifying with age. Indeed, the only deregulated pathway common to the three ages studied is relative to immune response (interferon gamma response) and we can note that two more (interferon alpha response and complement) are also slightly deregulated with a p-adjusted between 0.05 and 0.25 in some conditions. At 8 months, activation of immune pathways is confirmed

Gene ID	Gene Name	Localization-Start	Localization-End	Cer. 6w	Cer. 4m	Cer. 8m	Cortex 6w	Cortex 4m	Cortex 8m	Hip. 6w	Hip. 4m	Hip. 8m
Bub1b	Mitotic checkpoint serine/threonine kinase	118428692	118472072	Down	Down	Down	Down	Down	Down	Down	Down	Down
Ivd	Isovaleryl coenzyme A dehydrogenase	118692435	118713390	Down	Down	Down	Down	Down	Down	Down	Down	Down
Knl1	Kinetochore scaffold 1	118877600	118935982	//	Up	Up	Up	Up	Up	Up	Up	Up
Gchfr	GTP cyclohydrolase I feedback regulator	118998254	119002871	Up	Up	Up	//	//	//	//	//	//
Exd1	Exonuclease 3'-5' domain containing 1	119346986	119378108	Up	Up	Up	Up	Up	Up	Up	Up	Up
Jmjd7	Jumonji domain containing 7	119857964	119863075	Down	Down	Down	//	Down	Down	Down	Down	Down
Gm28042	Predicted gene, 28042	119857974	119873514	Down	Down	Down	//	Down	//	Down	Down	Down
Pla2g4b	Phospholipase A2, group IVB (cytosolic)	119863368	119873511	Down	Down	Down	//	Down	//	Down	Down	Down
Sptbn5	Spectrin beta, non-erythrocytic 5	119871974	119916159	Down	Down	Down	//	//	//	//	//	//
Pla2g4e	Phospholipase A2, group IVE	119996893	120075816	Up	Up	Up	Up	Up	Up	Up	Up	Up
Gm13997	Predicted gene 13997	120075402	120076079	//	Up	Up	Up	Up	Up	//	Up	Up
Tgm7	Transglutaminase 7	120924046	120946877	Down	Down	Down	//	//	//	//	//	//
Lcmt2	Leucine carboxyl methyltransferase 2	120958788	120971134	Down	Down	Down	//	Down	//	//	Down	//
Strc	Stereocilin	121194209	121217649	Down	Down	Down	//	Down	//	Down	Down	Down
Ell3	Elongation factor RNA polymerase II-like 3	121269491	121274759	Down	Down	Down	Down	Down	Down	Down	Down	Down
Mfap1b	Microfibrillar-associated protein 1B	121290716	121304548	//	//	//	Down	Down	Down	Down	Down	//
Gm14018	WD repeat domain 76 pseudogene	121315913	121321993	//	Down	//	Down	Down	Down	//	Down	//
Eif3j1	Eukaryotic translation initiation factor 3, subunit J1	121859027	121887079	Down	Down	Down	Down	Down	Down	Down	Down	Down
Spg11	Spatacsin vesicle trafficking associated	121884001	121948867	Down	Down	//	Down	Down	Down	Down	Down	Down
Bloc1s6	Biogenesis of lysosomal organelles complex-1, subunit 6, pallidin	122580423	122591395	Up	Up	Up	Up	Up	Up	Up	Up	Up
Fam227b	Family with sequence similarity 227, member B	125825403	125993924	Up	Up	Up	Up	Up	Up	Up	Up	Up
Fgf7	Fibroblast growth factor 7	125876578	125933105	Up	Up	Up	//	Up	Up	//	Up	Up

Table 2. Common DEGs localized on chromosome 2. Double slash means the gene is not significantly differentially expressed in this condition.

Process Category	Pathways	Cer. 6w	Cer. 4m	Cer. 8m	Cortex 6w	Cortex 4m	Cortex 8m	Hip. 6w	Hip. 4m	Hip. 8m
Cellular component	HALLMARK_APICAL_JUNCTION		-1.46			-1.46			-1.51	-1.40
Development	HALLMARK_ADIPOGENESIS	1.55		1.45			1.28	-1.38	-1.47	
	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	1.73					1.57		-1.81	
	HALLMARK_MYOGENESIS	1.22	-1.65			-1.58		-1.52	-1.79	-1.84
DNA damage	HALLMARK_DNA_REPAIR	1.51						-1.70		
	HALLMARK_UV_RESPONSE_DN	-1.26	1.55			1.47		1.49		
	HALLMARK_UV_RESPONSE_UP	1.54						-1.29		
Immune	HALLMARK_ALLOGRAFT_REJECTION		1.57	2.17			1.59			1.52
	HALLMARK_COAGULATION	1.40		2.13			1.62		-1.49	
	HALLMARK_COMPLEMENT	1.31	1.29	2.26			1.29			
	HALLMARK_INTERFERON_ALPHA_RESPONSE	1.85	1.78	1.76			1.60	-1.30	-1.29	
	HALLMARK_INTERFERON_GAMMA_RESPONSE	1.51	1.68	1.80			1.65			
	HALLMARK_IL6_JAK_STAT3_SIGNALING			2.27			1.33		-1.40	
	HALLMARK_INFLAMMATORY_RESPONSE	1.27		2.11			1.70			
Metabolic	HALLMARK_FATTY_ACID_METABOLISM	1.77	1.25						1.57	1.24
	HALLMARK_GLYCOLYSIS	1.49					-1.24			
	HALLMARK_OXIDATIVE_PHOSPHORYLATION	1.92			-1.65	-1.37	-1.35	-2.13	2.03	1.32
	HALLMARK_XENOBIOTIC_METABOLISM	1.48		1.55				-1.57		
Pathway	HALLMARK_APOPTOSIS	1.59		1.81			1.33			
	HALLMARK_PROTEIN_SECRETION			2.13	-1.50	1.66			2.02	1.80
	HALLMARK_UNFOLDED_PROTEIN_RESPONSE	1.26	1.37						1.71	1.38
Proliferation	HALLMARK_E2F_TARGETS	1.22	1.60			1.73	1.38		1.98	1.90
	HALLMARK_G2M_CHECKPOINT		1.71			2.02	1.45	1.66	1.68	1.46
	HALLMARK_MYC_TARGETS_V1	1.861	1.84		-1.39	1.79	1.26	-1.29	2.30	2.36
	HALLMARK_MYC_TARGETS_V2	1.68								
	HALLMARK_MITOTIC_SPINDLE	-1.57						1.66		
Signaling	HALLMARK_ANDROGEN_RESPONSE		1.53			1.41	1.68		1.62	1.39
	HALLMARK_ESTROGEN_RESPONSE_LATE	1.60								
	HALLMARK_IL2_STAT5_SIGNALING			1.88			1.33			
	HALLMARK_KRAS_SIGNALING_UP			1.94			1.66			
	HALLMARK_KRAS_SIGNALING_DN		-1.47			-1.481			-1.72	-1.63
	HALLMARK_MTORC1_SIGNALING	1.54	1.65				1.49		2.01	1.54
	HALLMARK_TNFA_SIGNALING_VIA_NFKB			1.77			1.75			
	HALLMARK_WNT_BETA_CATENIN_SIGNALING		-1.72			-1.43			-1.43	-1.35

p.adjusted < 0.05
p.adjusted < 0.25

Table 3. Summary of GSEA results using MsigDB Hallmark database. The classification of pathways was carried out according to[40]. The pathways selected are those with a p-adjusted < 0.05 in at least one of the nine conditions studied. In each table cell, colored in blue for negative or in orange for positive, is written the NES (normalized Enrichment score). A dark color corresponds to a p-adjusted < 0.05 and a light color to a p-adjusted < 0.25.

Class	Subclass	Pathway	Cer. 6w	Cer. 4m	Cer. 8m	Cortex 6w	Cortex 4m	Cortex 8m	Hip. 6w	Hip. 4m	Hip. 8m
Cellular Processes	Cellular community - eukaryotes	KEGG_FOCAL_ADHESION				1.47				-1.69	
	Transport and catabolism	KEGG_LYSOSOME	1.6		2.26					-1.31	
		KEGG_PEROXISOME	1.41								1.87
Environmental Information Processing	Signal transduction	KEGG_PHOSPHATIDYLINOSITOL_SIGNALING_SYSTEM	-1.99								
		KEGG_NOTCH_SIGNALING_PATHWAY		-1.97				-1.75		-1.94	-1.94
		KEGG_CALCIUM_SIGNALING_PATHWAY	-1.40							-1.54	-1.46
	Signaling molecules and interaction	KEGG_ECM_RECEPTOR_INTERACTION				1.77				-2.44	-1.69
Genetic Information Processing	Folding, sorting and degradation	KEGG_PROTEASOME	2.02						-1.77	1.80	1.70
		KEGG_PROTEIN_EXPORT	1.90					1.69		2.01	1.66
		KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS								1.69	
	Transcription	KEGG_SPLICEOSOME	1.52					2.03		1.90	1.93
	Translation	KEGG_RIBOSOME	2.39				1.65		-2.00	1.73	2.07
Human Diseases	Cancer: specific types	KEGG_BASAL_CELL_CARCINOMA	-1.38	-1.82				-1.65		-1.59	-1.66
	Cardiovascular disease	KEGG_ARRHYTHMOGENIC_RIGHT_VENTRICULAR_CARDIOMYOPATHY_ARVC								1.88	
		KEGG_HYPERTROPHIC_CARDIOMYOPATHY_HCM								1.73	
	Endocrine and metabolic disease	KEGG_TYPE_II_DIABETES_MELLITUS	-1.96								
	Immune disease	KEGG_SYSTEMIC_LUPUS_ERYTHEMATOSUS					2.25			1.62	1.74
	Neurodegenerative disease	KEGG_PARKINSONS_DISEASE	2.03					-1.64	-2.26	1.93	
		KEGG_HUNTINGTONS_DISEASE	1.45				-1.43	-1.62	-1.71	1.52	
		KEGG_ALZHEIMERS_DISEASE						-1.64	-1.63	1.59	
Metabolism	Energy metabolism	KEGG_OXIDATIVE_PHOSPHORYLATION	2.13				-1.79	-1.72	-2.41	1.88	
	Nucleotide metabolism	KEGG_PYRIMIDINE_METABOLISM								-1.67	
		KEGG_PURINE_METABOLISM						-1.50	-1.55		
	Xenobiotics biodegradation and metabolism	KEGG_METABOLISM_OF_XENOBIOTICS_BY_CYTOCHROME_P450	2.05								
		KEGG_DRUG_METABOLISM_CYTOCHROME_P450	2.03								
Organismal Systems	Development and regeneration	KEGG_AXON_GUIDANCE								-1.55	
	Immune system	KEGG_COMPLEMENT_AND_COAGULATION_CASCADES	1.55		2.34					-1.57	

p.adjusted < 0.05
p.adjusted < 0.25

Table 4. Summary of GSEA results using KEGG database. The classification of pathways was carried out according to KEGG database. The pathways selected are those with a p-adjusted < 0.05 in at least one of the nine conditions studied. In each table cell, colored in blue for negative or in orange for positive, is written the NES (normalized Enrichment score). A dark color corresponds to a p-adjusted < 0.05 and a light color to a p-adjusted < 0.25.

with up-regulation of the seven pathways of the MsigDB database relative to the immune category (allograft rejection; coagulation; complement; interferon alpha response; interferon gamma response; IL6 JAK STAT3 signaling; inflammatory response). Moreover, the three KEGG pathways that appear deregulated in cerebellum at 8 months can be linked to the immune system: complement and coagulation cascades, systemic lupus erythematosus and lysosome. Interestingly, we see up regulation of immune response also in the cortex but later than in cerebellum, with an activation of all MsigDB immune pathway at 8 months (with a p-adjusted between 0.05 et 0.25). Probably linked to this inflammatory response[39], perturbation of oxidative phosphorylation is visible even if it is not directionally concordant in every condition. Immune response alterations were less important in hippocampus.

Other deregulated pathways involved in genetic information processing, particularly in hippocampus, were also altered. Indeed, at 4 months in hippocampus, KEGG pathways relative to proteasome, protein export, ubiquitin mediated proteolysis, spliceosome and ribosome are up regulated. These pathways are still activated at 8 months even if the significance is lower. Concerning the cerebellum and the cortex, the disruption of these pathways is still notable but seems more diffuse. In the cortex only the spliceosome and protein export pathways are deregulated at 4 months and the ribosome pathway at 6 weeks. In the cerebellum, the disturbances could only be highlighted at 6 weeks with up-regulation of the proteasome, ribosome, protein export and spliceosome pathways.

We also identified a new category of processes deregulated in $Spg11^{-/-}$ mice, indeed, at least one pathway belonging to the proliferation category in the MsigDB database is deregulated in all conditions except in cerebellum at 8 months. The deregulation is particularly important for three pathways, E2F targets, G2M checkpoint and MYC targets V1, at 4 months in the three nervous structures and at 8 months in cortex and hippocampus.

Finally, we can notice variations in expression of genes involved in pathways relative to neuronal development such as axon guidance in hippocampus at 4 months, or notch signaling pathway at 4 months in cerebellum, cortex and in hippocampus at 4 months and 8 months. Pathways relative to neurodegenerative diseases, as Parkinson's disease, Huntington's disease and Alzheimer's disease, are also deregulated in six of nine conditions.

Discussion

To investigate the deregulation of cellular pathways in $Spg11^{-/-}$ mice, we conducted a comprehensive transcriptomic analysis using RNAseq technology. This sequencing was performed at three critical ages (6 weeks, 4 months and 8 months) across three different nervous system structures (cerebellum, cortex and hippocampus). This approach enabled us to unravel the pattern of gene deregulation with respect to both age and structural variations.

Upon comparing the DEGs across the nine conditions, we observed a significant deregulation of genes neighboring the $Spg11$ locus, particularly at early stages. This phenomenon could result from interdependent gene regulations in the region and has been associated to the so-called "passenger effect"[35,36]. Interestingly, another study on $Gatm$, encoding for Agat enzyme and localized on chromosome 2 in mouse in the immediate vicinity of $Spg11$, showed similar effects on transcriptomics. Choe et al. have generated an AGAT-deficient mice by homologous recombination in mouse embryonic stem cells, using a clone 129/SvJ BAC that was injected into C57BL/6J mouse blastocysts[41], a generation process similar to that used to created the $Spg11^{-/-}$ mouse[10]. The transcriptomic analysis of the $AGAT^{-/-}$ mouse revealed 17 DEGs compared to WT mouse including 11 genes on chromosome 2 close to $Spg11$ and $Gatm$ loci. Moreover, 5 genes exhibited deregulation in the same direction as our transcriptomic data: $Pla2g4e$, $Exd1$, Ivd, $Lcmt2$ and $Bloc1s6$[42], further emphasizing the interconnectedness of these genes in chromatin regulation, potentially explained by the gene passenger effect. Given these findings, it is crucial to take into account this potential gene passenger expression when analyzing $Spg11$ mouse models. Investigation into whether the $Gatm$ and $Spg11$ gene regions contain an overlapping TAD (Topologically Associated Domain) influencing the expression of common deregulated genes is also warranted since we cannot rule out that the mutation might have affected the local gene expression regulation.

Thus, the involvement other neighboring genes in the $Spg11$ mouse phenotype cannot be excluded, especially since some of them have known functions in cellular processes identified as altered by the absence of spatacsin. For instance, $Bloc1s6$ encodes a subunit of BLOC-1 complex that participates in endosomal trafficking and dendritic formation in hippocampal neurons[43]. Expression variation of $Pla2g4e$ in hippocampal neurons was correlated with cognitive deficits in elderly APP/PS1 mice[44]. To mitigate potential confounding effects, we opted to exclude these genes from our global GSEA analyzes in order to minimize their impact identifying pathways involved in pathological mechanisms.

The results of PCA and differential analysis suggested that gene expression has a different pattern in cerebellum than in hippocampus and cortex. Notably, the number of DEGs between 4 months and 8 months is stable in cerebellum, in contrast to the cortex and hippocampus where number of DEGs at 8 months decreases. This decrease is difficult to explain, as the symptoms in the mouse model do not diminish at 8 months. One possible explanation is the reported decrease in neurons relative to astrocytes in the cortex between these ages. Since many dysregulated genes at 4 months are likely neuron-specific, a reduction in neuron numbers could lower overall mRNA levels, decreasing upregulated genes and increasing downregulated ones based on threshold selection. Another potential explanation is a compensatory mechanism that may reduce transcript dysregulation in the cortex and hippocampus. Regarding a significant portion of DEGs is associated with inflammation in the cerebellum at 8 months, and GSEA confirmed an up-regulation of immune system pathways. In cortex and hippocampus, the inflammatory pathways are less activated at this age. Therefore, we can hypothesize that the pathological mechanisms leading to neurodegeneration in the cerebellum differ from those in the cortex and hippocampus, at least for a certain length of time.

Moreover, both in cerebellum and cortex at 8 months we can notice six of dowregulated mitochondrial genes: *mt-Atp6*, *mt-Atp8*, *mt-Co2*, *mt-Co3*, *mt-Nd4l* and *mt-Ts1*. Five of these six genes encodes subunits of OXPHOS complexes. Neuro-inflammation and defects in oxidative phosphorylation are two interconnected processes. Indeed, the oxidative phosphorylation process plays a key role in the cell by providing ATP, but an inefficient mechanism may lead to reactive oxygen species (ROS) generation, leading to a release of cytochrome c into the cytosol and finally an activation of mitochondria-mediated apoptotic pathway[45]. These processes have been described in several neurodegenerative diseases such as Parkinson disease, Alzheimer disease or amyotrophic lateral sclerosis (ALS)[46–48]. Moreover, differential expression of oxidative phosphorylation genes has been observed in patients with AD[49] or ALS[50]. Thus, we can assume that a decrease expression of genes coding for subunits of OXPHOS complexes is correlated with neuronal death in *Spg11*[−/−] mice. In addition, our results suggest that neuroinflammation plays a potential role in our model, which was functionally demonstrated recently through microgliosis and T-lymphocyte recruitment in the *Spg11*[−/−] model used in our RNASeq[51] and is consistent with the astrogliosis observed in the cortex at 8 months in the previous characterization of the mouse model[10].

The *Spg11*[−/−] mouse model previously characterized by Branchu et al. demonstrated disruptions in the lysosome-autophagy machinery, a finding corroborated by various studies on different Spg11 models. More specifically, autofluorescent aggregates were observed in Lamp1-positive vesicles as well as defects in the ALR mechanism[10,14–17]. Additionally, the absence of spatacsin correlated with defects in neuronal development, including disruption of neurite outgrowth and branching[18–21,23,24]. Therefore, we expected gene expression perturbations in the cellular pathways regulating these processes.

At 4 months we can effectively note deregulated genes belongings to lysosome and autophagy pathways in the three brain structures studied, suggesting possible biological modification in these cellular processes. However, in the enrichment analyzes few pathways relating to lysosomes and autophagy are statistically significant. Thus, these analyzes allowed us to highlight other cellular functions or pathways disrupted by the absence of spatacsin that could be more relevant for understanding the pathological process at work in *Spg11*[−/−] mice.

Indeed, GO analysis of DEGs and GSEA based on KEGG database indicated a down regulation of neuronal linked processes, especially synapse and axons development. Thus, at 4 months we noticed in cortex and hippocampus down-regulated DEGs associated with post synapse and cellular projection GO term. Moreover, GSEA confirmed the down regulation of axon guidance pathway in hippocampus. This observation aligns with the observed defects in Spg11 models. Additionally, gene expression variations were more pronounced at 4 months and one common deregulated pathway at this age was the Notch signaling pathway. This pathway regulates essential cell functions including cell proliferation and differentiation in various tissues[52]. In nervous system, notch signaling plays a critical role in neuronal development and adult neurogenesis[53,54]. Moreover, notch signaling is involved in regulation of neurite development[55–57] and synapse plasticity[58,59] notably in hippocampus[60]. Interestingly, in neural models derived from SPG11 iPSCs and in SPG11 organoids, there was an observed increase in neuroblast proliferation and premature neurogenesis that could potentially be corrected by treating the cells or organoids with tideglusib, a GSK3β inhibitor[22,61]. However, since GSK3β modulates the Notch pathway, where its deletion in mice leads to increased Notch activity and consequent inhibition of neurogenesis[62], it would be intriguing to test tideglusib in the *Spg11*[−/−] model to assess its effects on the Notch pathway and, more broadly, on neuronal development.

Moreover, at 4 months in hippocampus, and to a lesser extent in the cortex, we observed deregulations in genetic information processing pathways such as up regulation in genes related to spliceosome, ribosome, or globally to RNA process. Many studies highlighted an important role for RNA dysregulation in several neurodegenerative diseases[63,64]. For example, a TDP-43 mutation linked to ALS induces splicing alterations contributing to the early manifestation of the disease[65]. Together, these observations could indicate a neurodegeneration process.

Our data further reveal that the loss of *Spg11* functions causes widespread transcriptomic changes with alteration in cellular proliferation. Specifically, we identified a significant increase in E2F targets and MYC targets pathways. The E2F transcription factors family is known to play a crucial role in cell cycle progression[66]. E2F target genes can be classified in five groups, genes involved in: cell cycle regulation, DNA replication, mitosis and mitotic spindle checkpoint, DNA repair and checkpoint controls[67]. Moreover, many of the molecules participating in the cell cycle in the proliferating cells are associated with the E2F apoptotic pathway in postmitotic neurons[68]. Regarding Myc target genes, they are involved in cell cycle regulatory, in protein biosynthesis and apoptosis[69–71]. Both, E2F and MYC, have been linked to neurodegenerative diseases[72,73]. Aberrant expressions of E2F-1 or MYC transcription factor have been observed in neuron of patients with Parkinson's disease[74], Alzheimer's disease[75,76] and ALS[77]. Consequently, DNA damage and re-entry in cell cycle has emerged as common pathogenic mechanism in neurological diseases. Furthermore, study on SPG11-NPCs has revealed dysregulation of genes expression involved in cell cycle and neurogenesis pathways, and defect in proliferation leading to a reduction of neuronal cells numbers[22]. The early deregulations of proliferation pathways, from 6 weeks, and in every structure means that they precede neuronal death in the *Spg11*[−/−] mouse since this could only be observed from 4 months. We can therefore make the hypothesis that an early disruption in these cellular pathways triggers a process leading to neuronal death in agreement with the developmental delay detected in cell models of Mishra et al., 2016.

Our study is limited by its reliance on bulk RNA-seq, which averages data across cell types, and its exclusive focus on RNA without addressing protein expression, reducing the ability to capture cell-specific and functional insights. Future approaches like single-nucleus RNA sequencing (snRNA-seq), spatial transcriptomics, and multiplex immunostaining on histological slides offer promising solutions to uncover cell-specific transcriptional changes, spatial gene expression patterns, and protein localization, providing a more detailed understanding of molecular mechanisms.

In summary, our experimental study sheds light on the transcriptomic changes resulting from the loss of spatacsin function. These findings underscore the dysregulation of various cellular pathways associated with

neurodegeneration, including neuroinflammation, neurite and synapse development, and RNA metabolism. The deregulations are particularly prominent at 4 months, suggesting strong neurodegeneration at this stage. Furthermore, early deregulations in cell proliferation regulation further suggest its involvement in the pathology's onset. These results open the way to functional studies to explore these highlighted alterations.

Data availability

Sequence data that support the findings of this study have been deposited in the SRA database with the primary accession code PRJNA1130700. Code available on GitHub: https://github.com/Daniel2601/TranscriptomicScripts.

References

1. Boutry, M., Morais, S. & Stevanin, G. Update on the Genetics of Spastic Paraplegias. *Curr. Neurol. Neurosci. Rep.* **19**, (2019).
2. Hedera, P. Hereditary spastic paraplegia overview. *GeneReviews ®* (1993).
3. Méreaux, J. L. et al. Clinical and genetic spectra of 1550 index patients with hereditary spastic paraplegia. *Brain* **145**, 1029–1037 (2022).
4. Stevanin, G. et al. Mutations in SPG11, encoding spatacsin, are a major cause of spastic paraplegia with thin corpus callosum. *Nat. Genet.* **39**, 366–372 (2007).
5. Pozner, T., Regensburger, M., Engelhorn, T., Winkler, J. & Winner, B. Janus-faced spatacsin (SPG11): involvement in neurodevelopment and multisystem neurodegeneration. *Brain* **143**, 2369–2379 (2020).
6. Stevanin, G. et al. Mutations in SPG11 are frequent in autosomal recessive spastic paraplegia with thin corpus callosum, cognitive decline and lower motor neuron degeneration. *Brain J. Neurol.* **131**, 772–784 (2008).
7. Manole, A. et al. Severe axonal neuropathy is a late manifestation of SPG11. *J. Neurol.* **263**, 2278–2286 (2016).
8. Cardozo-Hernández, A. L., de Rezende, C., França, M. C. & T. J. R. & Hereditary spastic paraplegia type 11 (SPG11) is associated with obesity and hypothalamic damage. *J. Neurol. Sci.* **416**, 116982 (2020).
9. Puech, B. et al. Kjellin Syndrome: long-term neuro-ophthalmologic follow-up and novel mutations in the SPG11 gene. *Ophthalmology* **118**, 564–573 (2011).
10. Branchu, J. et al. Loss of spatacsin function alters lysosomal lipid clearance leading to upper and lower motor neuron degeneration. *Neurobiol. Dis.* **102**, 21–37 (2017).
11. Orlacchio, A. et al. SPATACSIN mutations cause autosomal recessive juvenile amyotrophic lateral sclerosis. *Brain J. Neurol.* **133**, 591–598 (2010).
12. Denora, P. S. et al. Motor neuron degeneration in spastic paraplegia 11 mimics amyotrophic lateral sclerosis lesions. *Brain* **139**, 1723–1734 (2016).
13. Chen, Y. & Yu, L. Recent progress in autophagic lysosome reformation. *Traffic* **18**, 358–361 (2017).
14. Chang, J., Lee, S. & Blackstone, C. Spastic paraplegia proteins spastizin and spatacsin mediate autophagic lysosome reformation. *J. Clin. Invest.* **124**, 5249–5262 (2014).
15. Varga, R. E. et al. In vivo evidence for lysosome depletion and impaired autophagic clearance in Hereditary Spastic Paraplegia Type SPG11. *PLoS Genet.* **11**, e1005454 (2015).
16. Renvoisé, B. et al. Lysosomal abnormalities in hereditary spastic paraplegia types SPG15 and SPG11. *Ann. Clin. Transl Neurol.* **1**, 379–389 (2014).
17. Boutry, M. et al. Inhibition of Lysosome Membrane Recycling Causes Accumulation of Gangliosides that Contribute to Neurodegeneration. *Cell. Rep.* **23**, 3813–3826 (2018).
18. Güner, F. et al. Axon-specific mitochondrial Pathology in SPG11 Alpha Motor neurons. *Front. Neurosci.* **15**, (2021).
19. Chen, Z. et al. Inhibiting mitochondrial fission rescues degeneration in hereditary spastic paraplegia neurons. *Brain* **145**, 4016–4031 (2022).
20. Martin, E. et al. Spatacsin and spastizin act in the same pathway required for proper spinal motor neuron axon outgrowth in zebrafish. *Neurobiol. Dis.* **48**, 299–308 (2012).
21. Pérez-Brangulí, F. et al. Dysfunction of spatacsin leads to axonal pathology in SPG11-linked hereditary spastic paraplegia. *Hum. Mol. Genet.* **23**, 4859–4874 (2014).
22. Mishra, H. K. et al. GSK3ß-dependent dysregulation of neurodevelopment in SPG11-patient iPSC model. *Ann. Neurol.* https://doi.org/10.1002/ana.24633 (2016).
23. Southgate, L. et al. Novel SPG11 mutations in Asian kindreds and disruption of spatacsin function in the zebrafish. *Neurogenetics* **11**, 379–389 (2010).
24. Pozner, T. et al. Tideglusib rescues neurite pathology of SPG11 iPSC derived cortical neurons. *Front. Neurosci.* **12**, (2018).
25. Liao, Y., Smyth, G. K. & Shi, W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**, 923–930 (2014).
26. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**, 139–140 (2010).
27. Robinson, M. D. & Oshlack, A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
28. Wickham, H. *Ggplot2: Elegant Graphics for Data Analysis* (Springer Cham, 2016).
29. Gao, C. H., Yu, G., Cai, P. & ggVennDiagram An intuitive, easy-to-Use, and highly customizable R package to generate venn diagram. *Front. Genet.* **12**, 706907 (2021).
30. Lawrence, M. et al. Software for computing and annotating genomic ranges. *PLoS Comput. Biol.* **9**, (2013).
31. Hahne, F. & Ivanek, R. Visualizing genomic data using Gviz and Bioconductor. *Methods Mol. Biol. Clifton NJ.* **1418**, 335–351 (2016).
32. Van der Maaten, L. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, (2008).
33. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R Package for comparing biological themes among Gene clusters. *OMICS J. Integr. Biol.* **16**, 284–287 (2012).
34. Liberzon, A. et al. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* **27**, 1739–1740 (2011).
35. Lusis, A. J., Yu, J. & Wang, S. S. The problem of passenger genes in transgenic mice. *Arterioscler. Thromb. Vasc Biol.* **27**, 2100–2103 (2007).
36. Berghe, T. V. et al. Passenger mutations confound interpretation of all genetically modified congenic mice. *Immunity* **43**, 200–209 (2015).
37. Subramanian, A. et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U S A.* **102**, 15545–15550 (2005).
38. Kanehisa, M. & Goto, S. K. E. G. G. Kyoto Encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30 (2000).

39. Park, J. et al. Mitochondrial ROS govern the LPS-induced pro-inflammatory response in microglia cells by regulating MAPK and NF-κB pathways. *Neurosci. Lett.* **584**, 191–196 (2015).
40. Liberzon, A. et al. The Molecular signatures database (MSigDB) hallmark gene set collection. *Cell. Syst.* **1**, 417–425 (2015).
41. Choe, C. et al. l-arginine:glycine amidinotransferase deficiency protects from metabolic syndrome. *Hum. Mol. Genet.* **22**, 110–123 (2013).
42. Jensen, M. et al. Homoarginine- and creatine-dependent gene regulation in murine brains with l-Arginine:glycine amidinotransferase deficiency. *Int. J. Mol. Sci.* **21**, (2020).
43. Ito, A. et al. Pallidin is a novel interacting protein for cytohesin-2 and regulates the early endosomal pathway and dendritic formation in neurons. *J. Neurochem.* **147**, 153–177 (2018).
44. Pérez-González, M. et al. PLA2G4E, a candidate gene for resilience in Alzheimer´s disease and a new target for dementia treatment. *Prog Neurobiol.* **191**, 101818 (2020).
45. Redza-Dutordoir, M. & Averill-Bates, Diana, A. Activation of apoptosis signalling pathways by reactive oxygen species. *Biochim. Biophys. Acta BBA - Mol. Cell. Res.* **1863**, 2977–2992 (2016).
46. Zhao, Y. et al. ROS-Induced mtDNA release: The emerging messenger for communication between neurons and innate immune cells during neurodegenerative disorder progression. *Antioxidants* **10**, (2021).
47. Obrador, E. et al. The link between oxidative stress, redox status, Bioenergetics and mitochondria in the pathophysiology of ALS. *Int. J. Mol. Sci.* **22**, (2021).
48. Singh, A., Kukreti, R., Saso, L. & Kukreti, S. Oxidative stress: a key modulator in neurodegenerative diseases. *Molecules* **24**, (2019).
49. Manczak, M., Park, B. S., Jung, Y. & Reddy, P. H. Differential expression of oxidative phosphorylation genes in patients with Alzheimer's disease: implications for early mitochondrial dysfunction and oxidative damage. *NeuroMolecular Med.* **5**, 147–162 (2004).
50. Ladd, A. C., Keeney, P. M., Govind, M. M. & Bennett, J. James P. mitochondrial oxidative phosphorylation transcriptome alterations in human amyotrophic lateral sclerosis spinal cord and blood. *NeuroMolecular Med.* **16**, 714–726 (2014).
51. Hörner, M. et al. CNS-associated T-lymphocytes in a mouse model of Hereditary Spastic Paraplegia type 11 (SPG11) are therapeutic targets for established immunomodulators. *Exp. Neurol.* **355**, 114119 (2022).
52. Zhou, B. et al. Notch signaling pathway: architecture, disease, and therapeutics. *Signal. Transduct. Target. Ther.* **7**, (2022).
53. Chen, N. et al. Effect of Notch1 on neural tube defects and neural stem cell differentiation induced by all–trans retinoic acid. *Mol. Med. Rep.* **23**, 1–1 (2021).
54. Ables, J. L., Breunig, J. J., Eisch, A. J. & Rakic, P. Not(Ch) just development: notch signalling in the adult brain. *Nat. Rev. Neurosci.* **12**, 269–283 (2011).
55. Redmond, L., Oh, S. R., Hicks, C., Weinmaster, G. & Ghosh, A. Nuclear Notch1 signaling and the regulation of dendritic development. *Nat. Neurosci.* **3**, 30–40 (2000).
56. Bonini, S. A. et al. Nuclear factor κB-Dependent neurite remodeling is mediated by Notch Pathway. *J. Neurosci.* **31**, 11697–11705 (2011).
57. Muroyama, Y., Baba, A., Kitagawa, M. & Saito, T. Olfactory sensory neurons control dendritic complexity of mitral cells via Notch Signaling. *PLoS Genet.* **12**, (2016).
58. Feng, S. et al. Notch1 deficiency in postnatal neural progenitor cells in the dentate gyrus leads to emotional and cognitive impairment. *FASEB J.* **31**, 4347–4358 (2017).
59. de Bivort, B. L., Guo, H. F. & Zhong, Y. Notch signaling is required for activity-dependent synaptic plasticity at the Drosophila neuromuscular junction. *J. Neurogenet.* **23**, 395–404 (2009).
60. Alberi, L. et al. Activity-induced notch signaling in neurons requires Arc/Arg3.1 and is essential for synaptic plasticity in hippocampal networks. *Neuron* **69**, 437–444 (2011).
61. Pérez-Branguli, F. et al. Human SPG11 cerebral organoids reveal cortical neurogenesis impairment. *Hum. Mol. Genet.* **28**, 961–971 (2019).
62. Kim, W. Y. et al. GSK-3 is a master regulator of neural progenitor homeostasis. *Nat. Neurosci.* **12**, 1390–1397 (2009).
63. Wilson, D. M. et al. Hallmarks of neurodegenerative diseases. *Cell* **186**, 693–714 (2023).
64. Chatterjee, B., Shen, C. K. J. & Majumder, P. R. N. A. Modifications and RNA metabolism in neurological Disease Pathogenesis. *Int. J. Mol. Sci.* **22**, (2021).
65. Arnold, E. S. et al. ALS-linked TDP-43 mutations produce aberrant RNA splicing and adult-onset motor neuron disease without aggregation or loss of nuclear TDP-43. *Proc. Natl. Acad. Sci. U. S. A.* **110**, E736–E745 (2013).
66. Dyson, N. The regulation of E2F by pRB-family proteins. *Genes Dev.* **12**, 2245–2262 (1998).
67. Ren, B. et al. E2F integrates cell cycle progression with DNA repair, replication, and G2/M checkpoints. *Genes Dev.* **16**, 245–256 (2002).
68. Greene, L. A., Biswas, S. C. & Liu, D. X. Cell cycle molecules and vertebrate neuron death: E2F at the hub. *Cell. Death Differ.* **11**, 49–60 (2004).
69. Adhikary, S. & Eilers, M. Transcriptional regulation and transformation by myc proteins. *Nat. Rev. Mol. Cell. Biol.* **6**, 635–645 (2005).
70. Dang, C. V. c-Myc target genes involved in cell growth, apoptosis, and metabolism. *Mol. Cell. Biol.* **19**, 1–11 (1999).
71. Grandori, C., Cowley, S. M., James, L. P. & Eisenman, R. N. The Myc/Max/Mad Network and the Transcriptional Control of Cell Behavior. *Annu. Rev. Cell. Dev. Biol.* **16**, 653–699 (2000).
72. Verdaguer, E., de Arriba Susana, G., Clemens, A., Pallàs, M. & Camins, A. Implication of the transcription factor E2F-1 in the modulation of neuronal apoptosis. *Biomed. Pharmacother.* **61**, 390–399 (2007).
73. Marinkovic, T. & Marinkovic, D. Obscure involvement of MYC in neurodegenerative diseases and neuronal repair. *Mol. Neurobiol.* **58**, 4169–4177 (2021).
74. Höglinger, G. U. et al. The pRb/E2F cell-cycle pathway mediates cell death in Parkinson's disease. *Proc. Natl. Acad. Sci. U. S. A.* **104**, 3585–3590 (2007).
75. Lee, H. P., Kudo, W., Zhu, X., Smith, M. A. & Lee, H. Early induction of c-Myc is Associated with neuronal cell death. *Neurosci. Lett.* **505**, 124–127 (2011).
76. Lee, H. et al. The neuronal expression of MYC causes a neurodegenerative phenotype in a Novel Transgenic mouse. *Am. J. Pathol.* **174**, 891–897 (2009).
77. Ranganathan, S., Scudiere, S. & Bowser, R. Hyperphosphorylation of the retinoblastoma gene product and altered subcellular distribution of E2F-1 during Alzheimer's disease and amyotrophic lateral sclerosis. *J. Alzheimers Dis.* **3**, 377–385 (2001).

Acknowledgements

Author contributions

D.S. and L.T.M. carried out the trancriptomics analysis. L.T.M. wrote the first version of the manuscript, which was discussed by all authors. N.A., M.L., T.E., J.B., G.S. and K.E.H. collected and prepared the samples. G.S. supervised the project.

Declarations

Competing interests
The authors declare no competing interests.

Ethics
This work was performed on RNA from mice according to the agreement of the Ethic Committee "Charles Darwin" n° Ce5/2012/045 and mouse handling was performed by trained biologists as required by the law.

Additional information
Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-86337-9.

Correspondence and requests for materials should be addressed to G.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.