Proceedings of the 3rd International Conference on Modern Medicine and Global Health
DOI: 10.54254/2753-8818/2025.21460

Identifying Differential Gene Expression Related to Spinocerebellar Ataxia Type 3 in the Brain of Model Mice to Reveal Potential Pathological Mechanisms

Zhihan Chen[1,a,*]

[1]*Keystone Academy, Anfu Street No.10, Beijing, China*
a. zhihan.chen@student.keystoneacademy.cn
**corresponding author*

Abstract: Spinocerebellar ataxia type 3 (SCA3) is an autosomal dominantly inherited neurodegenerative disorder caused by aggregation of polyglutamine ataxin-3 protein in the cerebellum and brainstem. The causative mechanisms of this disease have not been clearly identified in previous research, especially molecular changes that occur in the affected regions of the brain in disease phenotypes. Different differential gene expression (DGE) events identified in the brain of SCA3 mice can provide insights into biological changes occurring through stages of disease pathogenesis. RNA sequencing (RNA-seq) data collected from the brainstem and cerebellum of SCA3 model mice and wild-type (WT) mice of different ages is studied with DGE analysis programs to determine differentially expressed genes (DEGs) at different times in the lives of disease model mice compared to WT mice. No DEG was identified in the brain of 2-month-old SCA3 model mice, and 18, 10, and 5 DEGs were identified in the brain of 4-, 12-, and 17.5-month-old SCA3 model mice, respectively. The genes are involved in various distinct biological processes, including neurotransmission, white matter maturation, cell signaling, cell metabolism, and more. Different DEGs are identified in the brain of 4-, 12-, and 17.5-month-old SCA3 model mice, causing biological changes that may lead to disease phenotypes in early, mid, and late stages of disease progression. Neuronal dysfunction is a major consequence of differential expression in 4- and 12-month-old mice while DGE has more extended effects in 17.5-month-old mice.

Keywords: spinocerebellar ataxia type 3, model mice, differentially expressed genes.

1. Introduction

Spinocerebellar ataxia type 3 (SCA3), or Machado-Joseph disease (MJD), is a type of spinocerebellar ataxia characterized by the progressive loss of movement coordination and balance. Patients may also display other symptoms, including dystonia (muscle tensing), spasticity (muscle stiffness), tremors, dysarthria (difficulties in articulating speech), and oculomotor abnormalities[1]. SCA3 is believed to be the leading cause of ataxia worldwide, with an overall prevalence rate of 1 to 5 individuals in 100,000[2]. It is the most common type of dominant spinocerebellar ataxia, accounting for 59.6% of all SCA cases in Brazil[3] and 62.6% in China[4].

SCA3 is a genetic disease inherited in an autosomal dominant manner, caused by a mutation in the ATXN3 gene causing variably sized CAG trinucleotide repeat expansions in the 10th exon, ranging

DOI: 10.54254/2753-8818/2025.21460

from 60~87 repeats compared to 12~43 repeats in normal alleles[5]. The repeat expansions cause ataxin-3 protein, the translated product, to have an abnormally long polyglutamine (polyQ) section. The mutant protein abnormally aggregates in the brain stem and cerebellum via liquid-liquid phase separation, causing neuronal dysfunction and loss[6]. Specifically, neuronal loss occurs for afferent nerve fibers from and nuclei related to the spine, vestibular system, and pons, leading to an enlarged fourth ventricle[7].

Currently, no effective cure has been developed for SCA3, and only supportive management is available for patients to delay progressive symptoms[8]. One potential approach towards future drug development is studying genetic regulators of expression of the ataxin-3 protein to determine genes that can be up-regulated or down-regulated to reduce toxicity of the mutant protein.

However, the molecular mechanisms that lead to mutant ataxin-3 pathogenicity have not been fully elucidated. Suggested mechanisms include nuclear localization of aggregates of the mutant protein[9], impaired proteasome function[10], mitochondrial damage[11], and transcriptional changes[12]. Several different transcriptional dysregulation events have been shown in the brain of SCA3 model mice by some previous research: Ramani et al. (2017) [13] found that *Acy3* and *Tnfrsf13c* are up-regulated in oligodendrocytes, which could be involved in a toxic gain-of-function mechanism; Schuster et al. (2022) [14] identified 43 dysregulated genes that impaired oligodendrocyte maturation; and Toonen et al. (2018) [12] showed dysregulation of *Tmc3*, *Zfp488*, *Car2*, and *Chdh*, which are involved in the adrenergic and CREB signaling pathways. Overall, results from previous research show that transcriptional dysregulation events in SCA3 may be related to oligodendrocyte function and cell signaling events; however, large discrepancies exist in the current literature since conclusions from similar research vary and no consensus has been established on the specific pathological mechanism of the disease based on available research.

Previous research that studied transcriptional changes in the brain of SCA3 model mice adhered to similar methods for data analysis, in which RNA-seq expression profiling data from high throughput sequencing is collected and analyzed with bioinformatics techniques for studying differential expression—namely edgeR[12], DESeq[13], and limma[14], which will be further described in the Materials and Methods section. Then, the identified differentially expressed transcripts are studied with protein analysis programs such as Ingenuity Pathway Analysis[15] to determine the potential biological functions and involvement in biological pathways of the identified genes to suggest potential pathological mechanisms.

Several potential causes in the method of data collection could account for the discrepancies in previous research as described before. Firstly, mouse models of SCA3 often do not exhibit effective disease phenotypes due to factors such as low expression levels of mutant protein, the difference between the lifespan of mice and humans, and the lack of neural complexity of mice compared to humans, which may lead to variations between different models. For this reason, an integrated analysis of different mouse models may be significant for SCA3 research using mouse models[13]. Additionally, differences in differential expression analysis programs may account for the discrepancies from similar research since the programs are based on different algorithms, which have specific contexts for optimal performance. Thus, obtaining overlapping results from the same set of data using different programs may reveal significant findings.

This study aims to compare gene expression data of SCA3 model mice from various datasets to find DGE patterns that occur during different times in the lives of SCA3 model mice using bioinformatics tools to determine the specific transcriptional dysregulation events that are involved at different points in time in SCA3 pathogenesis. It is hypothesized that different genes would be differentially expressed during different times of the lifetime of SCA3 model mice due to changes during the progression of the disease. Thus, by analyzing differentially expressed genes during different times in the lifetime of SCA3 mice, this study aims to elucidate potential involvement of

DOI: 10.54254/2753-8818/2025.21460

differentially expressed genes in pathological mechanisms of SCA3 and determine potential therapeutic sites.

2. Methodology

2.1. Data collection

2.1.1. Data selection

To study DGE in SCA3 model mice compared to WT mice at different ages, RNA-seq data from the following datasets from Gene Expression Omnibus (GEO) will be analyzed: GSE145613, GSE261670, and GSE107958. Information about these datasets is summarized in table 1. Two SCA3 mouse models are studied in these datasets: a novel ataxin-3 knock-in (KI) mouse model is used in GSE145613, which contains 304 CAG/CAA repeats in both murine ataxin-3 alleles[16], and Q84, or MJD84.2, transgenic mouse model is used in GSE261670 and GSE107958 and the transgene contains 84 CAG repeats in exon 10 of the human ATXN3 gene[17].

Table 1: Overview of information of datasets used in this study. All datasets collect data from the cerebellum and most from the brainstem.

GEO Accession	Genotype of mouse model	Sample size for each genotype	Brain tissue of sample collection	Age of mice
GSE145613	WT	5 (for each age)	Cerebellum	2 and 12 months
	Ataxin-3 knock-in	5 (for each age)		
GSE261670	WT	4-6 (for each age)	Brainstem	4 months
	Q84 transgenic	4-6 (for each age)		
GSE107958	WT	8	Brainstem and cerebellum	17.5 months
	MJD84.2 transgenic	6		

RNA-seq data generated from samples collected in the brainstem and cerebellum will be used, since mutant ataxin-3 is found to cause the most significant pathological changes in these brain regions. To study differentially expressed genes in SCA3 model mice in early, mid, and late stages of the mice's lifetime, samples collected from mice aged 2, 4, 12, and 17.5 months will be analyzed separately.

2.1.2. Data pre-processing

RNA-seq data is organized in the form of an expression matrix in which rows indicate gene names and columns indicate sample name (a convention required by DGE analysis programs). For clearing data, duplicate values for gene names and null values for expression levels are removed to not interfere with further data processing. Group information (i.e., WT or SCA3) is assigned to each sample by aligning the expression matrix and a matrix containing group information for each sample. Then, normalization of data is performed. Data is normalized using the trimmed mean of M-values (TMM) method for analysis using edgeR[18]. TMM estimates scale factors from samples by trimming away extreme log-fold changes via a trimmed mean, or a mean produced from removing a fixed upper and lower percentage of the data. It can provide a more accurate representation of gene expression levels and adjust for differences in library sizes between samples[19]. DESeq2 and limma internally normalize data during analysis.

DOI: 10.54254/2753-8818/2025.21460

2.2. Differential gene expression analysis

The DGE analysis programs edgeR, DESeq2, and limma are used to generate three lists of DEGs. These methods are used to obtain separate DEG lists since different statistical methods usually yield different lists. The Benjamini-Hochberg (BH) method[20] is then used for p-value adjustment; genes with adjusted p-value < 0.05 are selected for DGE analysis.

Then, DGE analysis of the pre-processed data is performed using Generalized linear model (GLM) likelihood ratio test in edgeR, GLM fit with negative binomial distribution in DESeq2, and linear model fit in limma, all by contrasting data for SCA3 model mice against WT mice. In all 3 programs, genes with adjusted p-value > 0.05 are considered unchanged, genes with log 2 fold-change < -1 are considered to be down-regulated, and genes with log 2 fold-change > 1 are considered to be up-regulated. The 3 programs produce 3 DEG lists; then, DEGs common to all three gene lists obtained from the three methods are selected as a gene list.

2.3. Biological analysis

To gain biological insights into the selected DEGs, gene enrichment analysis is performed with Database for Annotation, Visualization and Integrated Discovery (DAVID). It produces functional annotations for a list of genes, showing information for gene ontology (GO) and biological pathways of the enriched annotation terms, as well as providing functional annotation clustering to manifest the most common annotations across a gene list[21]. Thus, the biological functions and pathways commonly represented by DEGs will be elucidated, which will allow potential pathological mechanisms to be suggested. Details about the functional annotations determined by DAVID will be further researched via the Gene Ontology database[22].

3. Results

3.1. DGE analysis of mice aged 2 months

Raw read counts data of samples collected from cerebellum of 5 WT and 5 knock-in SCA3 model mice aged 2 months (from GEO accession: GSE145613) were analyzed. The raw counts data was first pre-processed by transforming into CPM data.

Analysis with edgeR, DESeq2, and limma all determined no DEGs. This demonstrates that no transcriptional dysregulation occurs during an early time in the development of SCA3 model mice compared to WT mice of the same age, which is likely due to lack of disease phenotype during early stages of development.

3.2. DGE analysis of mice aged 4 months

Transcripts Per Million (TPM) data of samples collected from brainstem of 5 WT and 4 Q84 transgenic model mice aged 4 months (from GEO accession: GSE261670) were analyzed. The effective counts data and TPM data generated from kallisto were imported and pre-processed by calculating gene-level offset to correct for differences in average transcript length between samples[23].

Analysis with edgeR yields 29 DEGs, analysis with DESeq2 yields 107 DEGs, and analysis with limma yields 19 DEGs. Ultimately, 18 DEGs are commonly selected by all three programs. One of the DEGs is the human *ATXN3* gene and is thus removed from the list, resulting in a final gene list of 17 genes. Expression levels of the 5 up-regulated genes are shown in figure 1(a). As shown in the figure, expression levels of these genes all show significant up-regulation in the brainstem of 4-month-old Q84 transgenic SCA3 model mice compared to WT mice. Figure 1(b) shows expression

DOI: 10.54254/2753-8818/2025.21460

levels of the 12 down-regulated genes. Expression levels of these genes all show significant up-regulation in the brainstem of 4-month-old Q84 transgenic SCA3 model mice compared to WT mice, with *Car2* having the highest expression level in WT mice and the most significant down-regulation in SCA3 model mice among all DEGs.

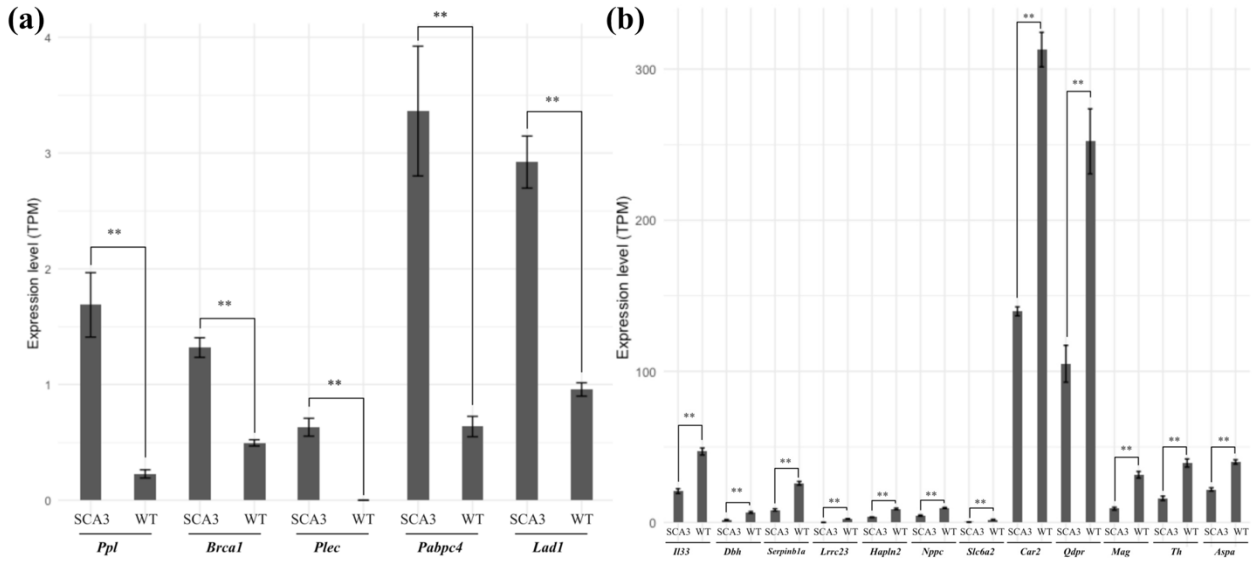

Figure 1: (a) Bar plot of expression level of up-regulated DEGs in brainstem of 4-month-old Q84 transgenic SCA3 model mice and WT mice. (b) Bar plot of expression level of down-regulated DEGs in brainstem of 4-month-old Q84 transgenic SCA3 model mice and WT mice. Bars show mean expression level (in TPM) across all samples and error bars show standard error of the mean. "**" indicates statistical significance at the level of adjusted p-value <0.05.

Results of gene enrichment analysis of the list of DEGs using DAVID are summarized in table 2. As shown in the table, several DEGs have biological functions that contribute to maintaining the normal functions of neurons, such as catecholamine biosynthesis, central nervous system myelination, intermediate filament cytoskeleton organization, transmission of nerve impulses, and folate biosynthesis. Cellular response to mechanical stimulus has the lowest p-value. Notably, two down-regulated genes positively regulate oligodendrocyte function, suggesting that SCA3 pathogenesis involves white matter changes.

Table 2: Functional annotations of DEGs obtained from "Functional Annotation Chart" from DAVID. Terms with p-value < 0.05 are selected and important result is highlighted.

Gene names	Term (s)	p-value
Il33, Mag, Plec	cellular response to mechanical stimulus	0.00142
Dbh, Th	catecholamine biosynthesis	0.00356
	tyrosine metabolism	0.0444
Mag, Aspa	central nervous system myelination	0.0115
Ppl, Plec	intermediate filament cytoskeleton organization	0.0130
Dbh, Slc6a2	response to pain	0.0130
Il33, Aspa	positive regulation of oligodendrocyte differentiation	0.0174
Mag, Plec	transmission of nerve impulse	0.0179
Qdpr, Th	folate biosynthesis	0.0291

DOI: 10.54254/2753-8818/2025.21460

3.3. DGE analysis of mice aged 12 months

Raw read counts data of samples collected from cerebellum of 5 WT and 5 knock-in SCA3 model mice aged 12 months (from GEO accession: GSE145613) were analyzed. The raw counts data was first pre-processed by transforming into CPM data.

Analysis with edgeR yields 22 DEGs, analysis with DESeq2 yields 17 DEGs, and analysis with limma yields 161 DEGs. Ultimately, 10 DEGs are commonly selected by all three programs. Figure 2(a) shows expression levels of the 2 up-regulated genes, demonstrating that significant up-regulation of genes occurs in the cerebellum of 12-month-old SCA3 model mice compared to WT mice. Expression levels of the 8 down-regulated genes are shown in figure 2(b). As shown in the figure, expression levels of these genes all show significant down-regulation in the cerebellum of 12-month-old SCA3 model mice compared to WT mice, with expression level of *Igfbp5* showing the most significant change among all down-regulated genes.

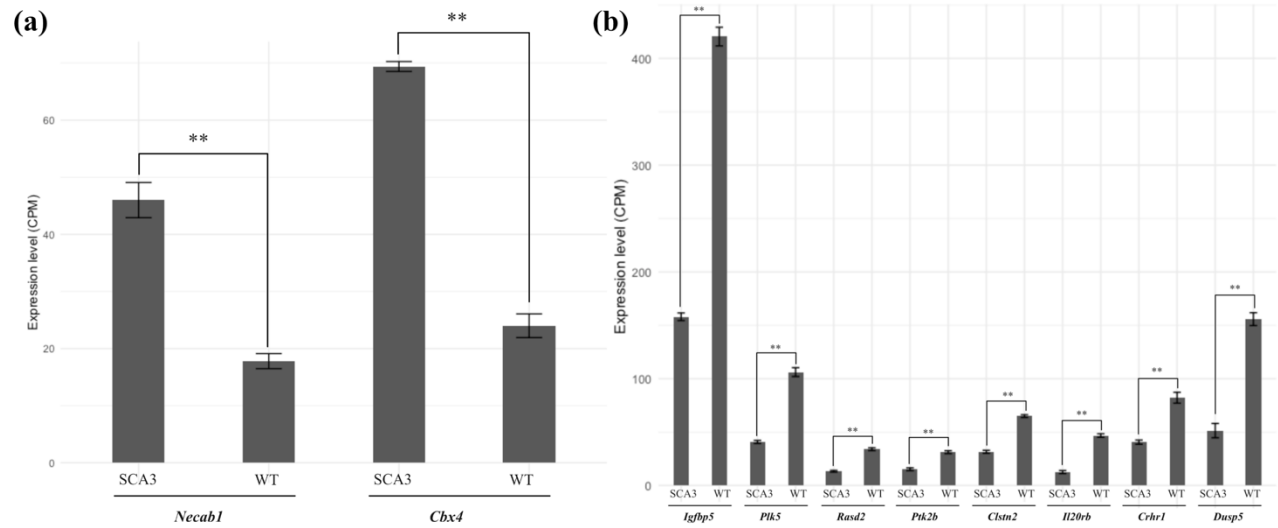

Figure 2: (a) Bar plot of expression level of up-regulated DEGs in cerebellum of 12-month-old SCA3 model mice and WT mice. (b) Bar plot of expression level of down-regulated DEGs in cerebellum of 12-month-old SCA3 model mice and WT mice.

Gene enrichment analysis of the list of genes as shown in table 3 using DAVID reveal several findings. From table 3, it can be observed that down-regulated genes are involved in several pathways that are critical for the maintenance of healthy neurons, including dendrite formation, positive regulation of cytosolic calcium ion concentration, negative regulation of neuron apoptotic process, long-term potentiation (LTP), and positive regulation of neuron projection development.

Table 3: Functional annotations of DEGs obtained from "Functional Annotation Chart" from DAVID. Terms with p-value < 0.05 are selected and important result is highlighted.

Gene names	Term (s)	p-value
Ptk2b, Rasd2, Igfbp5	positive regulation of protein kinase B signaling	0.00109
Clstn2, Ptk2b, Crhr1	long-term synaptic potentiation	0.0482
	positive regulation of cytosolic calcium ion concentration	0.0428
	negative regulation of neuron apoptotic process	0.0466
Ptk2b, Crhr1	response to hypoxia	0.0497
	cell surface receptor signaling pathway	0.0484
Ptk2b, Plk5	positive regulation of neuron projection development	0.0432

DOI: 10.54254/2753-8818/2025.21460

Additionally, a one-tailed ANOVA test was performed to determine if gene expression is affected by age in both SCA3 and WT mice using data from GSE145613, to ensure that the change in age from 2 months to 12 months for the same type of mouse model is not a factor that affects DGE . For WT mice, one-tailed ANOVA test determines that age is not correlated with gene expression with p-value > 0.05. However, for SCA3 mice, the same test determines that age is correlated with gene expression with p-value < 0.05.

3.4. DGE analysis of mice aged 17.5 months

Counts-per-million (CPM) data of samples collected from brainstem of 8 WT and 6 transgenic SCA3 model mice aged 17.5 months in GEO accession "GSE107958" were analyzed.

Analysis with edgeR yields 12 DEGs, analysis with DESeq2 yields 15 DEGs, and analysis with limma yields 18 DEGs. Ultimately, 5 DEGs are commonly selected by all three programs. As shown in figure 3, the expression levels of all DEGs in the brainstem of SCA3 model mice show significant up-regulation compared to their expression levels in the brainstem of WT mice.

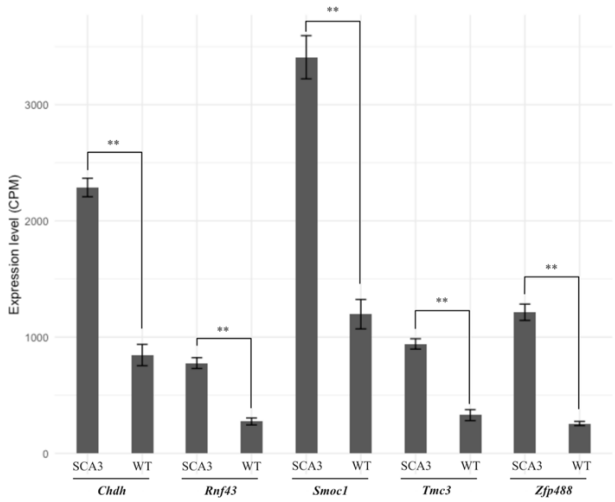

Figure 3: Bar plot of expression level of the DEGs in brainstem of 17.5-month-old SCA3 model mice and WT mice.

Gene enrichment analysis of the list of genes as shown in table 4 using DAVID reveal no shared functional annotations between 2 or more DEGs. The key biological functions of DEGs as determined by DAVID are summarized in table 4.

Table 4: Functional annotations of DEGs obtained from "Functional Annotation Chart" from DAVID. Terms with p-value < 0.05 are selected and important result is highlighted.

Gene names	Functional annotation
Chdh	mitochondrial autophagy; glycine, serine and threonine metabolism
Rnf43	Wnt signalling pathway; zinc-finger protein transmembrane domain
Smoc1	cell differentiation
Tmc3	ion transmembrane transport
Zfp488	positive regulation of oligodendrocyte differentiation; neurogenesis; positive regulation of myelination

DOI: 10.54254/2753-8818/2025.21460

4. Discussions

Different patterns in transcriptional dysregulation are observed in early, mid, and late times in the lifetime of SCA3 mice, revealing changes in disease progression over time.

No DEG was determined for SCA3 model mice with age of 2 months. This may be due to low expression levels of the disease protein in early developmental stages of SCA3 model mice, which is the case for KI disease model as shown by Ramani et al. (2017) [13]. Thus, transcriptional dysregulation does not occur during early stages of SCA3 pathogenesis, which agrees with the conclusion by Wiatr et al. (2019) [24].

DEGs identified in the brainstem of 4-month-old Q84 transgenic SCA3 model mice show that transcriptional changes begin to occur and hence disease phenotype begins to develop at this age. The identified DEGs partake in various biological processes, with many of them being responsible for normal neuronal functions. Differential expression of *Mag* and *Aspa* may cause changes in central nervous system myelination and differential expression of *Mag* and *Plec* may disrupt the transmission of nerve impulses, which would both impede the transmission of electric signals in the neuron, agreeing with clinical findings that neuronal dysfunction is a major SCA3 disease phenotype[25]. Moreover, transcriptional changes might also lead to changes in intermediate filament cytoskeleton organization and folate biosynthesis, which can further affect healthy neuronal function since intermediate filament organization is essential for the structural rearrangement and signaling regulation in neurons[26] and folate plays a role in central nervous system repair[27]. Additionally, differential expression of *Dbh* and *Th* may lead to changes in catecholamine biosynthesis, which may be correlated with the development of depression in SCA3 patients, which is a known SCA3 disease symptom[28]. Also, transcriptional dysregulation of *Dbh* and *Slc6a2* affects response to pain, which could potentially explain for feelings of pain in SCA3 patients[28]. Notably, down-regulation of *Il33* and *Aspa* would reduce oligodendrocyte differentiation, which agrees with the findings by Ramani et al. (2017) [13] and Schuster et al. (2022) [15] that impaired oligodendrocyte maturation is caused by early transcriptional changes in the brain of SCA3 model mice, which leads to white matter dysfunction via a toxic gain-of-function mechanism. The genes identified in this study, *Il33* and *Aspa*, were not identified in the previous studies and thus are of interest of research. Moreover, differential expression also affects cellular response to mechanical stimuli and tyrosine metabolism; however, how these biological changes relate to the development of SCA3 phenotype would require further investigation.

Several DEGs were identified in the cerebellum of 12-month-old SCA3 model mice, revealing transcriptional dysregulation events that contribute to pathogenesis during this stage of disease progression. Down-regulation of several genes that have normal functions related to processes that contribute to maintaining healthy neurons are observed, suggesting that impaired neuronal functions and potential neuronal loss likely occur during this stage of the disease. Down-regulation of *Ptk2b, Crhr1,* and *Clstn2* may reduce cytosolic calcium ion concentration, enhance neuron apoptotic process, disrupt the formation of LTP, and impair response to hypoxia. Additionally, down-regulation of *Ptk2b* and *Plk5* may reduce development of neuron projection. These biological changes in the neuron caused by differential expression of specific genes all interfere with normal neuronal functions, eventually leading to neuronal loss and death as observed in the SCA3 disease phenotype. Additionally, down-regulation of *Ptk2b, Rasd2,* and *Igfbp5* may lead to reduced protein kinase B signaling, causing impairment to cell-to-cell signaling.

Moreover, transcriptional changes in the brainstem and cerebellum of 17.5-month-old SCA3 model mice compare to WT mice reveal pathological changes that are different from those observed in SCA3 model mice of earlier ages. Interestingly, up-regulation of transcription is more common in SCA3 model mice of 17.5 months of age compared to those of 12 months of age, showing differences

DOI: 10.54254/2753-8818/2025.21460

in transcriptional dysregulation patterns during different stages in disease progression. Up-regulation of *Chdh* causes up-regulation of mitochondrial autophagy, which supports the findings by Yu et al. (2009) [11] that mitochondrial damage is a SCA3 disease phenotype. Up-regulation of *Rnf43* and *Smoc1* would cause changes to Wnt signalling pathway and cell differentiation, and the specific effects of differential expression of these two genes on disease phenotype require further research. Up-regulation of *Tmc3* may disrupt transport of ions across cell membranes of neurons, causing changes in intracellular and extracellular ion concentration that can impede the formation of electrical signals in neurons. *Zfp488,* which is up-regulated in the brainstem of 17.5 month old SCA3 model mice compared to WT mice, is responsible for positive regulation of oligodendrocyte differentiation, neurogenesis, and DNA methylation; thus, transcriptional dysregulation of this genes could lead to changes in white matter formation and gene expression. Up-regulation of *Zfp488* up-regulates oligodendrocyte maturation compared to in 4-month-old mice when down-regulated genes down-regulates oligodendrocyte maturation; further research is required to determine why this change in white matter formation occurs as the mice age.

Certain limitations exist for this study. The datasets used in this study have relatively low sample sizes, which causes each of them to have low statistical power, which is manifested in how sometimes results obtained using different DGE analysis programs from the same set of data disagree with each other and that some programs may produce an abnormally large list of DEGs. The SCA3 mouse models differ slightly from each other in that they are produced with different methods (i.e., transgenic or knock-in) and the number of CAG expansion repeats in the *ATXN3* gene differ slightly in different mouse models, which could lead to slight variations in the expression of the disease protein or the extent of neurodegeneration of disease phenotype. Additionally, result from the ANOVA test for 2 and 12 months old SCA3 mice unexpectedly reveals that gene expression is correlated with age, indicating that age is a confounding variable which diminishes the validity of DGE analysis results.

5. Conclusion

In conclusion, transcriptional dysregulation during early and middle stages of disease progression in SCA3 model mice mainly leads to impairment of neuronal functions and potential neuronal loss and death, while differential expression during late stage of the disease doesn't directly affect neuron functions and have more extended effects in the brain instead. For future research, white matter changes that occur over time in mice and human with SCA3 disease and the role of cell signaling and metabolic pathways in SCA3 pathogenesis can be investigated to further elucidate processes involved in SCA3 pathogenesis. Moreover, DEGs identified in this study can provide potential therapeutic possibilities. For example, certain disease phenotypes may be ameliorated by reducing the expression level of some up-regulated genes via RNA interference[29] or by increasing the expression level of some down-regulated genes via inhibition of mRNA-destabilizing factors or RNA interference pathways[30].

References

[1] *McLoughlin, H. S., Moore, L. R., & Paulson, H. L. (2020). Pathogenesis of SCA3 and implications for other polyglutamine diseases. Neurobiology of disease, 134, 104635. https://doi.org/10.1016/j.nbd.2019.104635*

[2] *Li, T., Martins, S., Peng, Y., Wang, P., Hou, X., Chen, Z., Wang, C., Tang, Z., Qiu, R., Chen, C., Hu, Z., Xia, K., Tang, B., Sequeiros, J., & Jiang, H. (2019). Is the High Frequency of Machado-Joseph Disease in China Due to New Mutational Origins? Frontiers in Genetics, 9. https://doi.org/10.3389/fgene.2018.00740*

[3] *de Castilhos, R. M., Furtado, G. V., Gheno, T. C., Schaeffer, P., Russo, A., Barsottini, O., et al. (2014). Spinocerebellar ataxias in Brazil–frequencies and modulating effects of related genes. Cerebellum 13, 17–28. doi: 10.1007/s12311-013-0510-y*

DOI: 10.54254/2753-8818/2025.21460

[4] Chen, Z., Wang, P., Wang, C., Peng, Y., Hou, X., Zhou, X., et al. (2018). Updated frequency analysis of spinocerebellar ataxia in China. Brain 141:e22. doi: 10.1093/brain/awy016

[5] Paulson H. (2012). Machado-Joseph disease/spinocerebellar ataxia type 3. Handbook of clinical neurology, 103, 437–449. https://doi.org/10.1016/B978-0-444-51892-7.00027-9

[6] Figueiredo, A. S., Loureiro, J. R., Macedo-Ribeiro, S., & Silveira, I. (2023). Advances in Nucleotide Repeat Expansion Diseases: Transcription Gets in Phase. Cells, 12(6), 826. https://doi.org/10.3390/cells12060826

[7] Paulson, H. L., Shakkottai, V. G., Clark, H. B., & Orr, H. T. (2017). Polyglutamine spinocerebellar ataxias - from genes to potential treatments. Nature reviews. Neuroscience, 18(10), 613–626. https://doi.org/10.1038/nrn.2017.92

[8] Paulson, H., & Shakkottai, V. (2020). Spinocerebellar Ataxia Type 3. Nih.gov; University of Washington, Seattle. https://www.ncbi.nlm.nih.gov/books/NBK1196/

[9] Paulson, H. L., Perez, M. K., Trottier, Y., Trojanowski, J. Q., Subramony, S. H., Das, S. S., Vig, P., Mandel, J. L., Fischbeck, K. H., & Pittman, R. N. (1997). Intranuclear inclusions of expanded polyglutamine protein in spinocerebellar ataxia type 3. Neuron, 19(2), 333–344. https://doi.org/10.1016/s0896-6273(00)80943-5

[10] Chai Y, Koppenhafer SL, Shoesmith SJ, Perez MK, Paulson HL. (1999) Evidence for proteasome involvement in polyglutamine disease: localization to nuclear inclusions in SCA3/MJD and suppression of polyglutamine aggregation in vitro. Hum Mol Genet. 8:673–82.

[11] Yu, Y. C., Kuo, C. L., Cheng, W. L., Liu, C. S., & Hsieh, M. (2009). Decreased antioxidant enzyme activity and increased mitochondrial DNA damage in cellular models of Machado-Joseph disease. Journal of neuroscience research, 87(8), 1884–1891. https://doi.org/10.1002/jnr.22011

[12] Toonen, L. J. A., Overzier, M., Evers, M. M., Leon, L. G., van der Zeeuw, S. A. J., Mei, H., Kielbasa, S. M., Goeman, J. J., Hettne, K. M., Magnusson, O. Th., Poirel, M., Seyer, A., 't Hoen, P. A. C., & van Roon-Mom, W. M. C. (2018). Transcriptional profiling and biomarker identification reveal tissue specific effects of expanded ataxin-3 in a spinocerebellar ataxia type 3 mouse model. Molecular Neurodegeneration, 13(1), 31. https://doi.org/10.1186/s13024-018-0261-9

[13] Ramani, B., Panwar, B., Moore, L. R., Wang, B., Huang, R., Guan, Y., & Paulson, H. L. (2017). Comparison of spinocerebellar ataxia type 3 mouse models identifies early gain-of-function, cell-autonomous transcriptional changes in oligodendrocytes. Human Molecular Genetics, 26(17), 3362–3374. https://doi.org/10.1093/hmg/ddx224

[14] Zeng, L., Zhang, D., McLoughlin, H. S., Zalon, A. J., Aravind, L., & Paulson, H. L. (2018). Loss of the Spinocerebellar Ataxia type 3 disease protein ATXN3 alters transcription of multiple signal transduction pathways. PloS one, 13(9), e0204438. https://doi.org/10.1371/journal.pone.0204438

[15] Schuster, K. H., Zalon, A. J., Zhang, H., DiFranco, D. M., Stec, N. R., Haque, Z., Blumenstein, K. G., Pierce, A. M., Guan, Y., Paulson, H. L., & McLoughlin, H. S. (2022). Impaired Oligodendrocyte Maturation Is an Early Feature in SCA3 Disease Pathogenesis. The Journal of Neuroscience, 42(8), 1604–1617. https://doi.org/10.1523/JNEUROSCI.1954-20.2021

[16] Haas, E., Incebacak, R. D., Hentrich, T., Huridou, C., Schmidt, T., Casadei, N., Maringer, Y., Bahl, C., Zimmermann, F., Mills, J. D., Aronica, E., Riess, O., Schulze-Hentrich, J. M., & Hübener-Schmid, J. (2022). A Novel SCA3 Knock-in Mouse Model Mimics the Human SCA3 Disease Phenotype Including Neuropathological, Behavioral, and Transcriptional Abnormalities Especially in Oligodendrocytes. Molecular neurobiology, 59(1), 495–522. https://doi.org/10.1007/s12035-021-02610-8

[17] Cemal, C. K., Carroll, C. J., Lawrence, L., Lowrie, M. B., Ruddle, P., Sahar Al-Mahdawi, King, R. H. M., Pook, M. A., Huxley, C., & Chamberlain, S. (2002). YAC transgenic mice carrying pathological alleles of the MJD1 locus exhibit a mild and slowly progressive cerebellar deficit. Human Molecular Genetics Online/Human Molecular Genetics, 11(9), 1075–1094. https://doi.org/10.1093/hmg/11.9.1075

[18] Smid, M., Coebergh van den Braak, R. R. J., van de Werken, H. J. G., van Riet, J., van Galen, A., de Weerd, V., van der Vlugt-Daane, M., Bril, S. I., Lalmahomed, Z. S., Kloosterman, W. P., Wilting, S. M., Foekens, J. A., IJzermans, J. N. M., MATCH study group, Martens, J. W. M., & Sieuwerts, A. M. (2018). Gene length corrected trimmed mean of M-values (GeTMM) processing of RNA-seq data performs similarly in intersample analyses while improving intrasample comparisons. BMC bioinformatics, 19(1), 236. https://doi.org/10.1186/s12859-018-2246-7

[19] Robinson, M. D., & Oshlack, A. (2010). A scaling normalization method for differential expression analysis of RNA-seq data. Genome biology, 11(3), R25. https://doi.org/10.1186/gb-2010-11-3-r25

[20] Benjamini Y and Hochberg Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol;57:289-300.

[21] Sherman, B. T., Hao, M., Qiu, J., Jiao, X., Baseler, M. W., Lane, H. C., Imamichi, T., & Chang, W. (2022). DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). Nucleic acids research, 50(W1), W216–W221. https://doi.org/10.1093/nar/gkac194

[22] Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J.

DOI: 10.54254/2753-8818/2025.21460

E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature genetics, 25(1), 25–29. https://doi.org/10.1038/75556

[23] Charlotte S, Michael I. L., Mark D. R. (2015). Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences. F1000Research http://dx.doi.org/10.12688/f1000research.7563.1

[24] Wiatr, K., Piasecki, P., Marczak, Ł., Wojciechowski, P., Kurkowiak, M., Płoski, R., Rydzanicz, M., Handschuh, L., Jungverdorben, J., Brüstle, O., Figlerowicz, M., & Figiel, M. (2019). Altered Levels of Proteins and Phosphoproteins, in the Absence of Early Causative Transcriptional Changes, Shape the Molecular Pathogenesis in the Brain of Young Presymptomatic Ki91 SCA3/MJD Mouse. Molecular Neurobiology, 56(12), 8168–8202. https://doi.org/10.1007/s12035-019-01643-4

[25] McCulloch, C. E. (2015). An Introduction to Generalized Linear Mixed Models. Conference on Applied Statistics in Agriculture. https://doi.org/10.4148/2475-7772.1314

[26] Bott, C. J., & Winckler, B. (2020). Intermediate filaments in developing neurons: Beyond structure. Cytoskeleton, 77(3-4), 110–128. https://doi.org/10.1002/cm.21597

[27] Iskandar, B. J., Nelson, A., Resnick, D., Skene, J. H., Gao, P., Johnson, C., Cook, T. D., & Hariharan, N. (2004). Folic acid supplementation enhances repair of the adult central nervous system. Annals of neurology, 56(2), 221–227. https://doi.org/10.1002/ana.20174

[28] Anelyssa D'Abreu, França, M. C., Paulson, H. L., & Iscia Lopes-Cendes. (2010). Caring for Machado–Joseph disease: Current understanding and how to help patients. Parkinsonism & Related Disorders (Online)/Parkinsonism & Related Disorders, 16(1), 2–7. https://doi.org/10.1016/j.parkreldis.2009.08.012

[29] Bumcrot, D., Manoharan, M., Koteliansky, V. et al. RNAi therapeutics: a potential new class of pharmaceutical drugs. Nat Chem Biol 2, 711–719 (2006). https://doi.org/10.1038/nchembio839

[30] Eberhardt, W., Doller, A., El-Sayed Akool, & Pfeilschifter, J. (2007). Modulation of mRNA stability as a novel therapeutic approach. Pharmacology & Therapeutics, 114(1), 56–73. https://doi.org/10.1016/j.pharmthera.2007.01.002

[31] The Gene Ontology Consortium. The Gene Ontology knowledgebase in 2023. Genetics. 2023 May 4;224(1):iyad031. DOI: 10.1093/genetics/iyad031

Appendix

This GitHub repository contains all the necessary information and files to reproduce results for this paper: https://github.com/zhihan2710/Pioneer-project.git.