



OPEN

## miRNA profiling as a complementary diagnostic tool for amyotrophic lateral sclerosis

Jack Cheng<sup>1,2,5</sup>, Wen-Kuang Ho<sup>1,5</sup>, Bor-Tsang Wu<sup>3</sup>, Hsin-Ping Liu<sup>4</sup>✉ & Wei-Yong Lin<sup>1,2</sup>✉

Amyotrophic lateral sclerosis (ALS), the most prevalent motor neuron disease characterized by its complex genetic structure, lacks a single diagnostic test capable of providing a conclusive diagnosis. In order to demonstrate the potential for genetic diagnosis and shed light on the pathogenic role of miRNAs in ALS, we developed an ALS diagnostic rule by training the model using 80% of a miRNA profiling dataset consisting of 253 ALS samples and 103 control samples. Subsequently, we validated the diagnostic rule using the remaining 20% of unseen samples. The diagnostic rule we developed includes miR-205-5p, miR-206, miR-376a-5p, miR-412-5p, miR-3927-3p, miR-4701-3p, miR-6763-5p, and miR-6801-3p. Remarkably, the rule achieved an 82% true positive rate and a 73% true negative rate when predicting the unseen samples. Furthermore, the identified miRNAs target 21 genes in the PI3K-Akt pathway and 27 genes in the ALS pathway, including notable genes such as BCL2, NEFH, and OPTN. We propose that miRNA profiling may serve as a complementary diagnostic tool to supplement the clinical presentation and aid in the early recognition of ALS.

ALS is a neurodegenerative disease that affects the central nervous system, characterized by motor neuron deficit and short life expectancy, but ALS can be challenging to diagnose, particularly in its early stages. Due to its rarity, physicians often consider more common illnesses before considering ALS, which can delay its diagnosis<sup>1</sup>. To improve early recognition, ALS diagnosis criteria have been proposed<sup>2</sup> and continuously revised<sup>3</sup>. However, even the recently proposed Gold Coast criteria<sup>4</sup> is still primarily based on clinical presentation, despite the fact that the genetic structure and biomarkers are gradually revealed by recent studies<sup>5</sup>. Additionally, newly developed predictive models, scales, and scoring systems can help patients and their physicians better understand the disease course. Although mechanism-based and potentially disease-modifying therapies are currently under clinical trials<sup>5</sup>, developing new diagnostic criteria and identifying genetic risk factors could also speed up the diagnostic process, and understanding the multisystem nature of ALS, including cognitive dysfunction and behavioral changes, is crucial for providing proper caregiving support and making end-of-life decisions.

ALS is currently divided into familial and sporadic categories. Familial ALS makes up 10–15% of cases and is inherited from family members with ALS or related syndromes like frontotemporal dementia<sup>6</sup>. About 70% of familial cases have mutations in known ALS genes. On the other hand, sporadic ALS makes up approximately 85% of cases and develops in patients without any family history of ALS. However, around 15% of sporadic ALS cases have private pathogenic mutations in known ALS genes, meaning they do not have a family history of ALS<sup>6</sup>. The cause of the remaining 85% of sporadic cases is unknown. Over 40 genes linked to ALS have been discovered, each varying in penetrance, frequency, and mode of inheritance. Among these genes, C9orf72, TARDBP, SOD1, and FUS are the most prevalent and have the highest penetrance<sup>7</sup>. Either toxic gain-of-function or loss-of-function mutations in these known genes are associated with ALS pathological processes. These mutations lead to protein aggregates forming, a key pathological feature in both sporadic and familial ALS cases<sup>8</sup>. The underlying pathophysiological processes can be broadly categorized into four main types: impaired RNA metabolism, altered proteostasis or autophagy, cytoskeletal or trafficking defects, and mitochondrial dysfunction<sup>9</sup>. For example, RNA metabolism is often affected by ALS-associated genes such as C9orf72, TARDBP, and FUS. Inclusions of TDP-43 and FUS can impair the normal function of DNA and RNA binding proteins, leading to significant changes in transcription and RNA processing. ALS-associated genes such as TDP-43 can also cause

<sup>1</sup>Graduate Institute of Integrated Medicine, College of Chinese Medicine, China Medical University, Taichung 40402, Taiwan. <sup>2</sup>Department of Medical Research, China Medical University Hospital, Taichung 40447, Taiwan. <sup>3</sup>Department of Senior Citizen Service Management, National Taichung University of Science and Technology, Taichung City 40343, Taiwan. <sup>4</sup>Graduate Institute of Acupuncture Science, College of Chinese Medicine, China Medical University, Taichung 40402, Taiwan. <sup>5</sup>These authors contributed equally: Jack Cheng and Wen-Kuang Ho. ✉email: hpliu@mail.cmu.edu.tw; linwy@mail.cmu.edu.tw

dysregulation of proteostasis and autophagy by preventing damaged protein clearance. Additionally, cytoskeletal and tubulin defects induced by mutant ALS genes like TUBA4A and PFN1 can interfere with axonal trafficking. Finally, SOD1 is a key gene linked to ALS and has been shown to cause mitochondrial dysfunction and increase oxidative stress, which is central to the development of the disease. Despite significant advances, our understanding of the complete molecular basis for the pathophysiology remains incomplete<sup>5</sup>.

MiRNAs are small non-coding molecules that post-transcriptionally regulate various biological processes, including neuronal function, development, and progression of diseases<sup>10</sup>. The expression level of miRNAs, including miR-1, miR-9, miR-133a, miR-133a/b, miR-142-3p, miR-149, miR-206, miR-223-3p, miR-326, miR-338-3p, miR-374b-5p, miR-424, and miR-451, in the serum, muscle, tissue, and cerebrospinal fluid of individuals with ALS differs from that of healthy controls<sup>11–17</sup>. These changes in miRNA expression may target the peripheral nerves, CNS, or skeletal muscles, potentially contributing to the etiology of ALS and playing a crucial role in its progression<sup>18</sup>. Therefore, Emde, et al., have proposed that different subtypes of ALS may share a common feature: the global dysregulation of miRNAs<sup>19</sup>. Thus, detecting changes in miRNA expression profile may serve as a valuable diagnostic biomarker to identify disease onset and progression. Additionally, identifying dysregulated miRNAs could offer a promising avenue for developing therapeutic approaches to treat ALS. However, due to ALS's polygenic nature and complexity, no single miRNA has been found universally dysregulated among previous studies. This fact makes it challenging to use miRNA as a diagnostic tool for ALS.

Here, we demonstrate the idea that ALS diagnosis could be assisted with the combination of the expression level of several circulating miRNAs. We established the judgement rule by machine learning using publically available dataset contributed by Magen and colleagues<sup>20</sup>, and showed that this rule model predicts 82% true positive (ALS) and 73% true negative (healthy control) in the blind dataset, and identified several novel miRNAs that target ALS genes.

## Methods

**Data source and batch correction.** The raw counts of miRNA were downloaded from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database<sup>21</sup> of the accession number GSE168714, contributed by Magen and colleagues<sup>20</sup>. The dataset was accessed at <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE168714> from March 1st to 3rd, 2023. The dataset comprises the annotated counts by small RNA-seq of the RNA extracted from the plasma of 253 ALS and 103 control<sup>20</sup>. The batch number of data collection and ALS/control state were also available with the raw count. To account for the differences in our study design compared to the original work of Magen and colleagues, we performed our own batch correction using the following method. Firstly, we temporarily excluded highly present miRNAs whose raw count represented more than 2% of the total counts (13 miRNAs excluded, as detailed in Supplementary File 1) during the calculation of batch correction coefficient. Secondly, we calculated the sum of miRNA counts for each sample. Thirdly, we determined the average sum of miRNA counts for each batch. We then defined a *batch coefficient* as the quotient of the *maximum average sum of all batches* divided by the *average sum of the batch*. Finally, we calculated the corrected count, including the temporarily excluded highly present miRNAs, by multiplying the raw counts by the *batch coefficient*.

**Machine learning.** The strategy to identify the key miRNAs was adopted from our previous studies<sup>22,23</sup>. Briefly, the transposition of the corrected miRNA counts served as the input file to the machine learning program RapidMiner<sup>24</sup>. For clarity, bold italicized text is used to denote the terminology in RapidMiner. In the RapidMiner process setting, patient ID was used as *ID*; patient ALS state was used as *Label*, and the corrected miRNA counts were used as *Attributes* to perform training. *Rule Induction* was adopted as the algorithm, and the overall training and validation program was shown in Supplementary File 2, with the following parameters. The *Split Data* operator separated the sample library into 80/20 sets by shuffled sampling for model building and independent validation, respectively. *Rule Induction* was performed with the criterion of *information gain*, *sample ratio* of 0.9, *pureness* of 0.9, and *minimal prune benefit* of 0.25, while ten times of *Cross Validation* was used to improve the model. Finally, the *Apply model* operator utilized the 20% blind set to validate the generated *Rule Model*, and the *Performance* operator demonstrated the performance of the model.

**mRNA targets of miRNAs.** DIANA-TarBase v7 was used to find the experimentally validated mRNA targets of miRNAs<sup>25</sup>. Data were accessed at <http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=tarbase/index> from April 13th to 20th, 2023.

**Interaction network and enrichment.** String-db version 11.5 was used to generate the interaction network and pathway enrichment<sup>26</sup>. Data were accessed at <https://string-db.org/> from April 20th to 23rd, 2023. Interaction network was generated with highest confidence (0.9) and disconnected nodes in the network were hidden.

**Intersection analysis.** Venny 2.0 was used to generate the Venn diagram<sup>27</sup>. Data was accessed at <https://bioinfogp.cnb.csic.es/tools/venny/index2.0.2.html> from April 20th to 23rd, 2023.

**Statistics.** Student's T-test was used to estimate the significance of the difference between two groups.

## Results

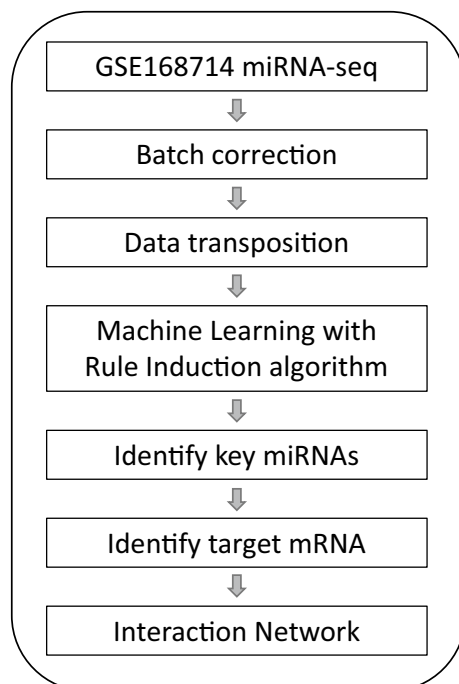
Machine learning has been proven to be a powerful tool to identify genetic biomarkers in neurodegenerative disorders with complex and heterogeneous genetic factors, such as Alzheimer's disease<sup>22</sup> and Huntington's disease<sup>23</sup>. Since ALS is also known for its complexity and heterogeneity genetic architecture, we applied machine learning on an ALS-control miRNA-seq dataset<sup>20</sup> to establish a judgment rule compromising miRNA profiles to identify ALS biomarkers following the strategy illustrated in Fig. 1. After batch correction and transposition, batch-corrected counts of miRNAs of 253 ALS and 103 control was generated (Supplementary File 3), and served as the input file for machine learning, where 80% samples served as the training set, and 20% as the unseen testing set. The generated model is shown in Fig. 2A, compromising miR-205-5p, miR-206, miR-376a-5p, miR-412-5p, miR-3927-3p, miR-4701-3p, miR-6763-5p, and miR-6801-3p, and its ROC curve in the training stage is shown in Fig. 2B. This rule model predicts 82% true positive and 73% true negative in the unseen dataset. The expression fold change of the identified miRNAs is shown in Fig. 3, with miR-412-5p, miR-3927-3p, miR-4701-3p, and miR-6801-3p significantly down-regulated in ALS, but not the other four miRNAs. This reflects the idea that a major advance of machine learning over traditional comparative methods in identifying biomarkers from expression profiles<sup>22</sup>.

To understand the role of miRNA-mediated biological processes in ALS pathology, we inquired the mRNA targets of miR-205-5p, miR-206, miR-376a-5p, miR-412-5p, miR-3927-3p, miR-4701-3p, miR-6763-5p, and miR-6801-3p in TarBase<sup>25</sup>, a curated database of miRNA-target interactions experimental evidence. Four miRNAs, i.e., miR-3927-3p, miR-4701-3p, miR-6763-5p, and miR-6801-3p, have no experiment-validated targets, while miR-205-5p, miR-206, miR-376a-5p, miR-412-5p have 296, 100, 43, and 200 targets, respectively (Supplementary File 4–7). Next, we showed that miRNA target genes enrich ALS and PI3K-Akt signaling pathways (Fig. 4, Supplementary File 8, 9), forming an interaction network with miR-205-5p targets occupying central hubs. Interestingly, The Venn diagram shows that the miRNA-target sets are nearly mutually exclusive in both ALS (Fig. 5A) and PI3K-Akt signaling pathway (Fig. 5B).

## Discussion

In this study, we proposed a diagnostic rule of ALS by the expression of miRNAs, including miR-205-5p, miR-206, miR-376a-5p, miR-412-5p, miR-3927-3p, miR-4701-3p, miR-6763-5p, and miR-6801-3p, where miR-206 has been reported in several previous ALS studies, while other miRNAs are novel in the ALS research field. Furthermore, these miRNAs mutually exclusively target genes in the ALS or PI3K-Akt pathways, which supports the idea that the combination of miRNAs, but not any single miRNA, may serve as a tool to facilitate ALS diagnosis. In the context of ALS's heterogeneous genetics, we discuss the identified miRNAs' and their target genes' relevance to ALS below.

**Heterogeneity of ALS genetics.** Recent advancements have greatly enhanced our comprehension of the genetic origins of familial ALS. Approximately 40–55% of familial ALS cases can be accounted for by variations in well-known ALS-linked genes<sup>7</sup>. While more than 50 potential causative or disease-modifying genes have been identified, pathogenic variants in SOD1, C9ORF72, FUS, and TARDBP are most frequent, whereas



**Figure 1.** The workflow of this study.

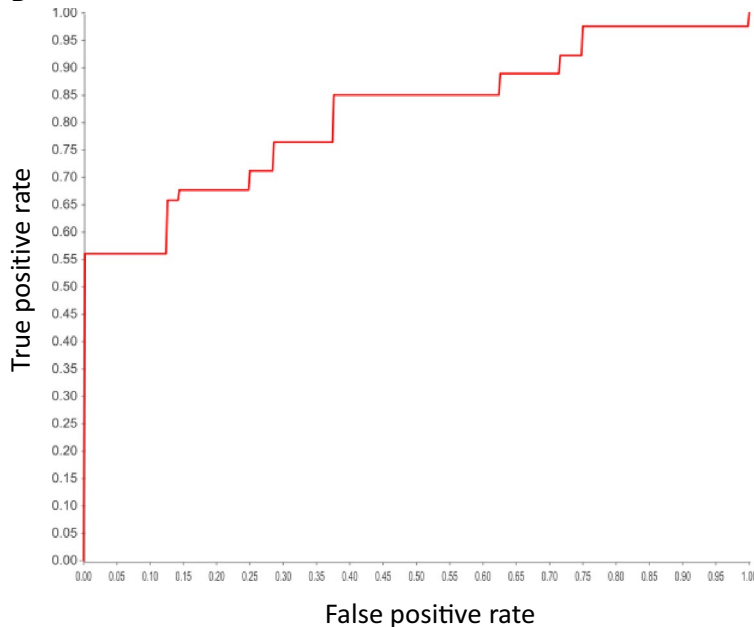
A

**Rule Model**

if hsa-miR-3927-3p  $\leq$  1.271 then 1 (2 / 120)  
 if hsa-miR-206 > 244.153 and hsa-miR-205-5p  $\leq$  107.280 and hsa-miR-4701-3p  $\leq$  37.774 then 1 (1 / 39)  
 if hsa-miR-376a-5p  $\leq$  26.846 and hsa-miR-412-5p > 20.492 then 0 (32 / 1)  
 if hsa-miR-6763-5p > 17.441 and hsa-miR-6801-3p  $\leq$  65.689 then 1 (2 / 23)  
 else 0 (36 / 22)

correct: 250 out of 278 training examples.

B



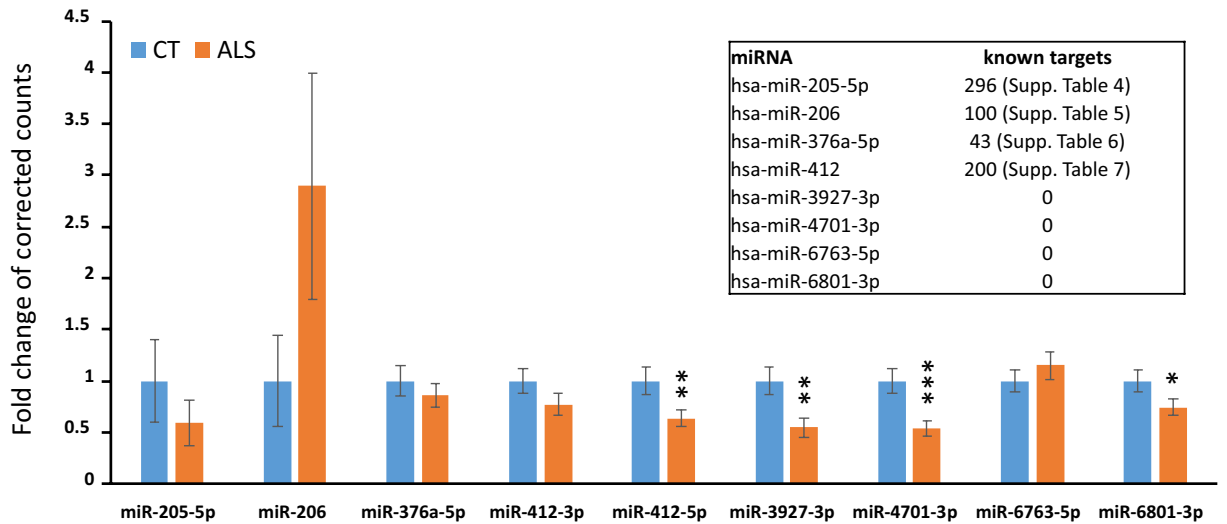
**Figure 2.** The predictive model generated by machine learning and its performance. (A) The rule model judging where a sample is ALS or control, where “1” denotes ALS and “0” denotes control. The value in the conditional expression is the batch-corrected count of the miRNA. (B) The ROC curve of the rule model, with AUC of 0.831.

variants in other genes are relatively uncommon<sup>28</sup>. However, in cases of sporadic ALS, diagnostic progress has only elucidated a fraction of the cases, with the etiology remaining unexplained in over 90% of patients<sup>29</sup>. It is widely accepted that genetic risk factors play a significant role in sporadic ALS, with heritability estimated at approximately 60% based on twin studies<sup>30</sup>. However, despite extensive genetic association studies, identifying heritable genetic risk factors in sporadic ALS remains elusive.

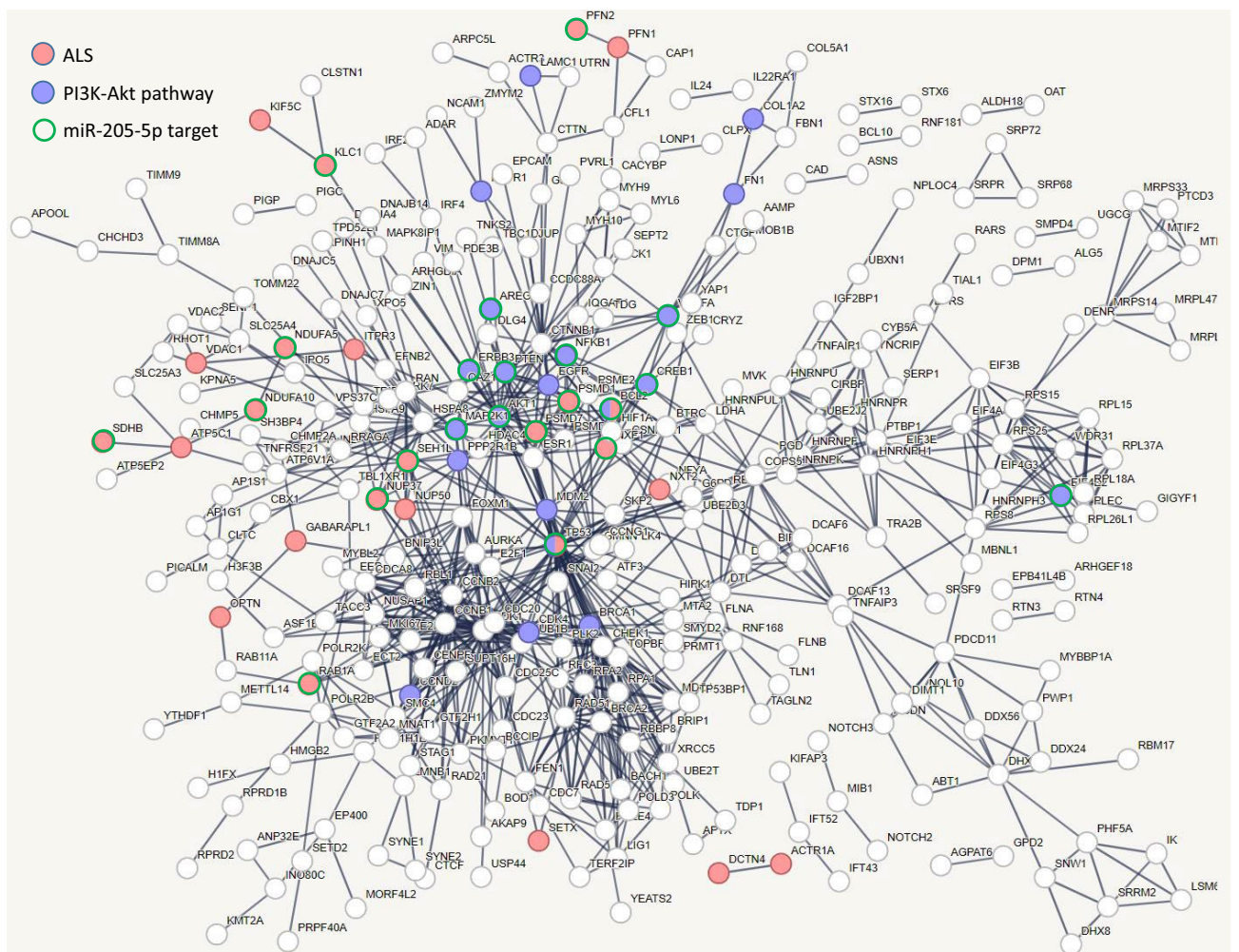
Despite decades of research, ALS’s underlying causative pathogenic mechanisms remain uncertain, particularly in sporadic cases. The development and progression of the disease are likely influenced by multiple factors rather than a single initiating event<sup>31</sup>. Moreover, genetic and phenotypic variations among patients pose challenges in comprehending and drawing conclusions about the general pathogenic mechanisms of ALS. Given the extensive number of genes and cellular processes implicated in ALS, numerous disease mechanisms have been proposed, including disruptions in RNA metabolism<sup>32</sup>, compromised protein homeostasis<sup>33</sup>, defects in nucleocytoplasmic transport<sup>34</sup>, impaired DNA repair<sup>35</sup>, mitochondrial dysfunction<sup>36</sup>, oxidative stress<sup>37</sup>, disturbances in axonal transport<sup>38</sup>, and oligodendrocyte dysfunction<sup>39</sup>. Further clarification is required to determine the timing and extent to which each of these mechanisms contributes to the pathogenesis of ALS.

As an attempt to see whether ALS clinical phenotypes could be differentiated by miRNAs, we constructed new cohorts of ALS patients from the same GSE168714 dataset by their clinical phenotypes: bulbar-onset or non-bulbar-onset, with 83 or 170 patients, respectively (Supplementary File 10), and these new cohorts of patients were subjected to de novo analysis in machine learning. However, the performance of the newly-established model was poor, with a recall rate of 75.9% and 24.6% for the cohorts, respectively. Thus, machine learning could not differentiate the cohorts of ALS, at least under the present condition. We further analyzed whether the identified miRNAs express differentially in the cohorts of ALS (Supplementary File 11). We found that miR-206 and miR-205 express differentially in the cohorts of ALS, but not significantly. We suspect that the sample size is

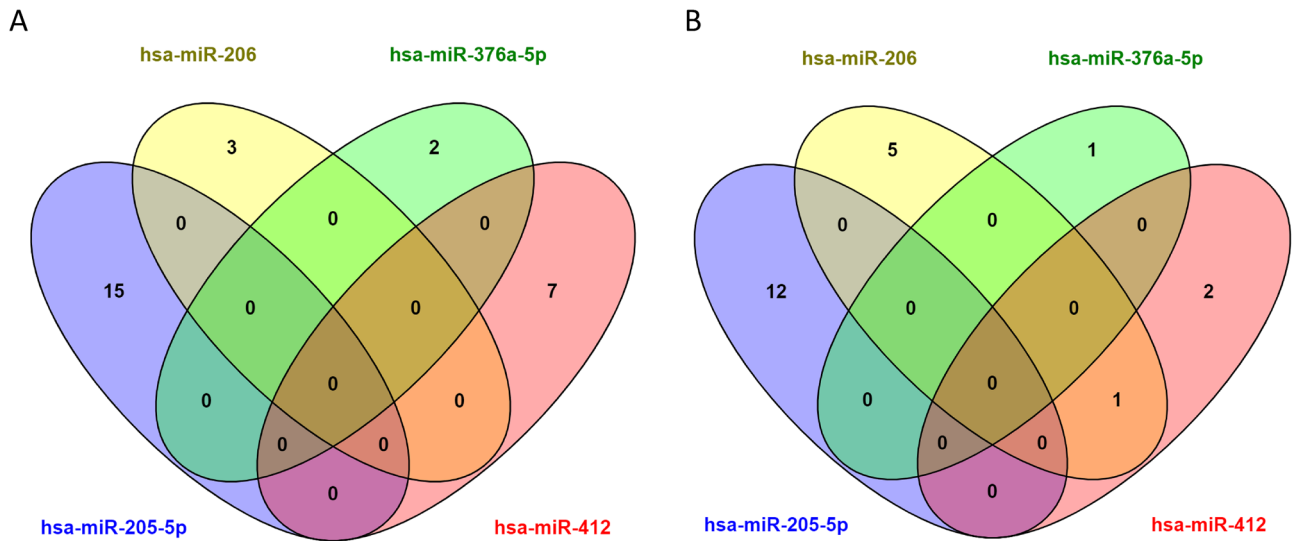




**Figure 3.** The miRNA expression fold change of the miRNAs identified by the rule model. Error bars stand for the standard error of the mean (SEM); while \*, \*\*, and \*\*\* stand for p-value < 0.05, 0.01, and 0.001 of Student's T-test, respectively. The upper right panel shows the number of experimentally validated mRNA targets of miRNA, with the full list provided in the corresponding supplementary files.



**Figure 4.** The protein interaction network of the mRNA targets of the identified miRNAs. Red balls indicate proteins in the Amyotrophic lateral sclerosis of KEGG, while blue balls indicate proteins in the PI3K-Akt signaling pathway of KEGG.



**Figure 5.** The Venn diagram showing that the miRNA-target sets are nearly mutually exclusive. (A) The ALS genes. (B) The PI3K-Akt genes.

the bottleneck to uncover the nature of the ALS cohorts. We also summarized other publicly-available miRNA datasets (Supplementary File 12). However, their sample size is insufficient for an independent machine learning study, and none of the datasets meets the requirement for validation of the current model that all eight miRNAs used in the machine learning model must be available.

**miR-206.** miR-206 participates in various stages of muscle differentiation, encompassing alternative splicing, DNA synthesis, and cell apoptosis<sup>40</sup>. During development, miR-206 hinders the activity of Pax7 and Pax3, effectively restricting the proliferative potential of satellite cells while promoting their differentiation into myogenic progenitor cells. Conversely, reducing miR-206 leads to the overexpression of Pax7 and Pax3, which consequently inhibits the differentiation of myoblasts. As Pax7 and Pax3 are known pro-survival factors, the down-regulation of miR-206 can induce apoptosis<sup>41</sup>. Thus, miR-206 plays a protective role and facilitates the regeneration of neuromuscular junctions following acute nerve injury, particularly in the context of ALS<sup>42</sup>. Knock-out of miR-206 delays and mutilates muscle reinnervation in ALS mouse models of SOD mutant<sup>43</sup>. Recent findings indicate elevated levels of miR-206 in the plasma of ALS patients and could indicate of disease progression<sup>44,45</sup>.

**miR-205.** miR-205 exhibits significant expression levels in various human epithelial tissues, including the breast, prostate, skin, eye, and thymus. Its primary role in these tissues is crucial to tissue morphogenesis and homeostasis. Specifically, it upholds the epithelial phenotype by directly targeting two transcription factors: zinc finger E-box-binding homeobox 1 (ZEB1) and ZEB2, which repress E-cadherin and other genes associated with polarity<sup>46</sup>. During the early stages of embryonic development, miR-205 is expressed in trophoblasts, where it regulates placental development by suppressing the Mediator of RNA polymerase II transcription subunit 1 (MED1)<sup>47</sup>. Moreover, in embryonic development, miR-205 governs the differentiation of extraembryonic endoderm and spermatogenesis by influencing cell migration and adhesion genes<sup>48</sup>. In the mammary gland, miR-205 displays high expression levels in the basal stem cells. Overexpression of miR-205 has been shown to induce the expansion of the progenitor cell population while reducing cell size and promoting cellular proliferation. These effects are achieved by repressing PTEN<sup>49</sup>. In this context, miR-205 regulates the production of the basement membrane protein complex laminin-332 and its receptor integrin- $\beta$ 4, thereby ensuring proper tissue polarity and morphogenesis<sup>50</sup>. In the skin epidermis and stratified epithelia of the esophagus and tongue, miR-205 has been found to play a significant role in expanding the stem cell population through its regulation of PI3K signaling<sup>51</sup>. Additionally, by influencing the same signaling pathways, miR-205 enhances the migration of human epidermal and corneal epithelial keratinocytes, thereby contributing to wound healing and corneal development<sup>52</sup>.

**miR-376a and miR-412.** The physiological roles of miR-376a and miR-412 are not fully understood yet, despite some studies reported its participation in cancer and neurological disorders. For example, increased levels of miR-376a have been observed in the T cells of patients with multiple sclerosis (MS)<sup>53</sup>, and miR-412 may inhibit clear cell renal cell carcinoma progression<sup>54</sup>. Conversely, in the late-onset form of Alzheimer's disease (LOAD), miR-376a has been identified as down-regulated in the brain<sup>55</sup>. Meanwhile, the expression change of miR-412 has been mentioned in the brain of alcohol use disorder<sup>56</sup>, and also in Alzheimer's disease<sup>57</sup>.

**BCL2.** BCL2 is targeted by miR-205<sup>58</sup> and controls caspase activation and the initiation of programmed cell death<sup>59</sup> and thus regulates neuronal development and neurodegeneration<sup>60</sup>. Several lines of evidence show that BCL2 probably involves in ALS pathological progression. Epidemiologic studies found altered expression of

BCL2 in ALS spinal cord motor neurons<sup>61</sup> and post-central gyrus<sup>62</sup>. In vivo studies showed that overexpression of BCL2 prolongs the survival of the ALS mouse model<sup>63</sup> and improves neuromuscular function<sup>64</sup>. In vitro studies revealed that ALS-associated mutant SOD1 aggregates BCL2<sup>65</sup> and advocates BCL2 conformational changes<sup>66</sup>.

**NEFH.** Neurofilament heavy polypeptide (NEFH) is one of the three intermediate filament proteins forming neurofilaments<sup>67</sup>. NEFH is targeted by miR-205<sup>68</sup> and could be phosphorylated by GSK3 $\beta$ <sup>69</sup> and regulate the Akt- $\beta$ -catenin pathway<sup>70</sup>. Moreover, epidemiological study showed that NEFH mutation<sup>71,72</sup> and expression<sup>73</sup> is associated with ALS. Besides, NEFH mutation or expression is associated with other disorders of central or peripheral neural system, including schizophrenia<sup>74</sup>, alcoholics<sup>75</sup>, and Charcot-Marie-Tooth neuropathy<sup>76</sup>.

**OPTN.** OPTN, also known as optineurin, is a highly conserved protein in various species<sup>77</sup>. OPTN is targeted by miR-205<sup>68</sup> as well. It plays diverse roles in vesicular trafficking, NF $\kappa$ B/NF- $\kappa$ B signaling, and autophagy. Specifically, OPTN has been identified as an autophagy receptor that facilitates the connection between ubiquitinated autophagy substrates and MAP1LC3/LC3-positive phagophore membranes<sup>78</sup>. Furthermore, mounting evidence suggests that OPTN acts as an inducer of autophagy, initiating the autophagic process<sup>79,80</sup>. Moreover, studies indicate that OPTN's involvement in autophagic initiation can commence as early as the formation of autophagosomal membranes<sup>81,82</sup>. These groundbreaking findings underscore the multifunctional role of OPTN as a potential autophagy receptor throughout the autophagic process, expanding beyond its traditional perception as a receptor operating solely at a single stage of autophagy. OPTN is gathering attention in ALS research, since variants of this gene are associated with ALS<sup>83–85</sup>. Moreover, OPTN mutant induces neuronal cell death by mediating mitophagy<sup>86</sup>, autophagy and ER stress<sup>87</sup>. Interestingly, OPTN mutation might be the common cause of ALS and corticobasal syndrome (CBS)<sup>88</sup>.

**PI3K in ALS.** The PI3K-akt signaling pathway governs metabolism, cell survival, motility, transcription, and cell-cycle progression. In recent years, studies have revealed the involvement of the PI3K-Akt signaling pathway in neurodegenerative diseases. For instance, butylphthalide has been shown to activate the PI3K-Akt/GSK-3 $\beta$  signaling pathway in an ischemic cerebral infarction model, reducing nerve function damage and protecting local nerve cells<sup>89</sup>. Therefore, therapeutic strategies for ALS targeting the PI3K-Akt pathway has been shown to increase anti-apoptotic protein expression levels, reduce pro-apoptotic protein expression levels, and improve cell survival rate and mitochondrial function in ALS<sup>90</sup>. Moreover, studies by Xiang and colleagues have found that AEG-1 can regulate the PI3K-Akt pathway<sup>91</sup>, and the absence of AEG-1 in ALS motor neurons inhibits the PI3K-Akt pathway and increases cell apoptosis<sup>91</sup>. Thus, dysregulated miRNAs may promote ALS pathology by mediating PI3K-Akt signaling pathway.

## Conclusion

In sum, we showed that a set of miRNA expressions could serve as a diagnostic tool for ALS, and these miRNAs target ALS and PI3K-Akt pathways in a mutually exclusive way. The key miRNAs include miR-205-5p, miR-206, and miR-376a-5p, while key targets are BCL2, NEFH, and OPTN. We propose that miRNA profiling may facilitate clinical presentation in the early recognition of ALS.

## Data availability

All data in this study are included in the supplementary data. The raw data used for machine learning is shown in Supplementary File 3, and the first two rows (file descriptions) must be removed before use. The raw counts of miRNA are also available from the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database of the accession number GSE168714.

## Code availability

The machine learning platform RapidMiner Studio is available at <https://rapidminer.com/>. The process is shown in Supplementary File 2, and the parameters are in the “Methods” section.

Received: 25 May 2023; Accepted: 17 August 2023

Published online: 23 August 2023

## References

- Ryan, M., Heverin, M., McLaughlin, R. L. & Hardiman, O. Lifetime risk and heritability of amyotrophic lateral sclerosis. *JAMA Neurol.* **76**, 1367–1374 (2019).
- Brooks, B. R. El Escorial World Federation of Neurology criteria for the diagnosis of amyotrophic lateral sclerosis. Subcommittee on Motor Neuron Diseases/Amyotrophic Lateral Sclerosis of the World Federation of Neurology Research Group on Neuromuscular Diseases and the El Escorial “Clinical limits of amyotrophic lateral sclerosis” workshop contributors. *J. Neurol. Sci.* **124**, 96–107 (1994).
- Brooks, B. R., Miller, R. G., Swash, M. & El Munsat, T. L. Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph. Lateral Scler. Other Motor Neuron Disord.* **1**, 293–299 (2000).
- Shefner, J. M. *et al.* A proposal for new diagnostic criteria for ALS. *Clin. Neurophysiol.* **131**, 1975–1978 (2020).
- Feldman, E. L. *et al.* Amyotrophic lateral sclerosis. *Lancet* **400**, 1363 (2022).
- Goutman, S. A. *et al.* Emerging insights into the complex genetics and pathophysiology of amyotrophic lateral sclerosis. *Lancet Neurol.* **21**, 465 (2022).
- Zou, Z.-Y. *et al.* Genetic epidemiology of amyotrophic lateral sclerosis: A systematic review and meta-analysis. *J. Neurol. Neurosurg. Psychiatry* **88**, 540–549 (2017).
- Mackenzie, I. R., Rademakers, R. & Neumann, M. TDP-43 and FUS in amyotrophic lateral sclerosis and frontotemporal dementia. *Lancet Neurol.* **9**, 995–1007 (2010).



9. Nguyen, H. P., Van Broeckhoven, C. & van der Zee, J. ALS genes in the genomic era and their implications for FTD. *Trends Genet.* **34**, 404–423 (2018).
10. Cai, Y., Yu, X., Hu, S. & Yu, J. A brief review on the mechanisms of miRNA regulation. *Genomics Proteomics Bioinform.* **7**, 147–154 (2009).
11. Russell, A. P. *et al.* Disruption of skeletal muscle mitochondrial network genes and miRNAs in amyotrophic lateral sclerosis. *Neurobiol. Dis.* **49**, 107–117 (2013).
12. De Felice, B. *et al.* miR-338-3p is over-expressed in blood, CFS, serum and spinal cord from sporadic amyotrophic lateral sclerosis patients. *Neurogenetics* **15**, 243–253 (2014).
13. Sumitha, R. *et al.* Differential expression of microRNA-206 in the gastrocnemius and biceps brachii in response to CSF from sporadic amyotrophic lateral sclerosis patients. *J. Neurol. Sci.* **345**, 254–256 (2014).
14. Hawley, Z. C., Campos-Melo, D. & Strong, M. J. Novel miR-b2122 regulates several ALS-related RNA-binding proteins. *Mol. Brain* **10**, 1–14 (2017).
15. Pegoraro, V., Merico, A. & Angelini, C. Micro-RNAs in ALS muscle: Differences in gender, age at onset and disease duration. *J. Neurol. Sci.* **380**, 58–63 (2017).
16. Saucier, D. *et al.* Identification of a circulating miRNA signature in extracellular vesicles collected from amyotrophic lateral sclerosis patients. *Brain Res.* **1708**, 100–108 (2019).
17. Banack, S. A., Dunlop, R. A., Stommel, E. W., Mehta, P. & Cox, P. A. miRNA extracted from extracellular vesicles is a robust biomarker of amyotrophic lateral sclerosis. *J. Neurol. Sci.* **442**, 120396 (2022).
18. Liu, H. *et al.* Systematic review and meta-analysis on microRNAs in amyotrophic lateral sclerosis. *Brain Res. Bull.* **194**, 83 (2023).
19. Emde, A. *et al.* Dysregulated miRNA biogenesis downstream of cellular stress and ALS-causing mutations: A new mechanism for ALS. *EMBO J.* **34**, 2633–2651 (2015).
20. Magen, I. *et al.* Circulating miR-181 is a prognostic biomarker for amyotrophic lateral sclerosis. *Nat. Neurosci.* **24**, 1534–1541 (2021).
21. Barrett, T. *et al.* NCBI GEO: Archive for functional genomics data sets—Update. *Nucleic Acids Res.* **41**, D991–D995 (2012).
22. Cheng, J., Liu, H.-P., Lin, W.-Y. & Tsai, F.-J. Machine learning compensates fold-change method and highlights oxidative phosphorylation in the brain transcriptome of Alzheimer's disease. *Sci. Rep.-UK* **11**, 13704 (2021).
23. Cheng, J., Liu, H.-P., Lin, W.-Y. & Tsai, F.-J. Identification of contributing genes of Huntington's disease by machine learning. *BMC Med. Genomics* **13**, 1–11 (2020).
24. Ristoski, P., Bizer, C. & Paulheim, H. Mining the web of linked data with rapidminer. *J. Web Semantics* **35**, 142–151 (2015).
25. Vlachos, I. S. *et al.* DIANA-TarBase v7.0: Indexing more than half a million experimentally supported miRNA: mRNA interactions. *Nucleic Acids Res.* **43**, D153–D159 (2015).
26. Szklarczyk, D. *et al.* The STRING database in 2021: Customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res.* **49**, D605–D612 (2021).
27. Oliveros, J. C. *Venny. An Interactive Tool for Comparing Lists with Venn's Diagrams.* (2007).
28. Boylan, K. Familial amyotrophic lateral sclerosis. *Neurol. Clin.* **33**, 807–830 (2015).
29. Renton, A. E., Chiò, A. & Traynor, B. J. State of play in amyotrophic lateral sclerosis genetics. *Nat. Neurosci.* **17**, 17–23 (2014).
30. Al-Chalabi, A. *et al.* An estimate of amyotrophic lateral sclerosis heritability using twin data. *J. Neurol. Neurosurg. Psychiatry* **81**, 1324–1326 (2010).
31. Mejzini, R. *et al.* ALS genetics, mechanisms, and therapeutics: Where are we now?. *Front. Neurosci.* **13**, 1310 (2019).
32. Mandrioli, J., Mediani, L., Alberti, S. & Carra, S. *Seminars in Cell & Developmental Biology.* 183–192 (Elsevier).
33. Thompson, A. G. *et al.* CSF extracellular vesicle proteomics demonstrates altered protein homeostasis in amyotrophic lateral sclerosis. *Clin. Proteomics* **17**, 1–12 (2020).
34. Lin, Y.-C. *et al.* Interactions between ALS-linked FUS and nucleoporins are associated with defects in the nucleocytoplasmic transport pathway. *Nat. Neurosci.* **24**, 1077–1088 (2021).
35. Lopez-Gonzalez, R. *et al.* Partial inhibition of the overactivated Ku80-dependent DNA repair pathway rescues neurodegeneration in C9ORF72-ALS/FTD. *Proc. Natl. Acad. Sci.* **116**, 9628–9633 (2019).
36. Cuzzolino, M. & Carri, M. T. Mitochondrial dysfunction in ALS. *Prog. Neurobiol.* **97**, 54–66 (2012).
37. Barber, S. C. & Shaw, P. J. Oxidative stress in ALS: Key role in motor neuron injury and therapeutic target. *Free Radical Biol. Med.* **48**, 629–641 (2010).
38. Bilsland, L. G. *et al.* Deficits in axonal transport precede ALS symptoms in vivo. *Proc. Natl. Acad. Sci.* **107**, 20523–20528 (2010).
39. Boillée, S. *et al.* Onset and progression in inherited ALS determined by motor neurons and microglia. *Science* **312**, 1389–1392 (2006).
40. Ma, G. *et al.* MiR-206, a key modulator of skeletal muscle development and disease. *Int. J. Biol. Sci.* **11**, 345 (2015).
41. Chen, J.-F. *et al.* microRNA-1 and microRNA-206 regulate skeletal muscle satellite cell proliferation and differentiation by repressing Pax7. *J. Cell Biol.* **190**, 867–879 (2010).
42. Williams, A. H. *et al.* MicroRNA-206 delays ALS progression and promotes regeneration of neuromuscular synapses in mice. *Science* **326**, 1549–1554 (2009).
43. De Felice, B. *et al.* A miRNA signature in leukocytes from sporadic amyotrophic lateral sclerosis. *Gene* **508**, 35–40 (2012).
44. de Andrade, H. M. *et al.* MicroRNAs-424 and 206 are potential prognostic markers in spinal onset amyotrophic lateral sclerosis. *J. Neurol. Sci.* **368**, 19–24 (2016).
45. Waller, R. *et al.* Serum miRNAs miR-206, 143-3p and 374b-5p as potential biomarkers for amyotrophic lateral sclerosis (ALS). *Neurobiol. Aging* **55**, 123–131 (2017).
46. Gregory, P. A. *et al.* The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat. Cell Biol.* **10**, 593–601 (2008).
47. Mouillet, J. F., Chu, T., Nelson, D. M., Mishima, T. & Sadovsky, Y. MiR-205 silences MED1 in hypoxic primary human trophoblasts. *FASEB J.* **24**, 2030–2039 (2010).
48. Li, C., Finkelstein, D. & Sherr, C. J. Arf tumor suppressor and miR-205 regulate cell adhesion and formation of extraembryonic endoderm from pluripotent stem cells. *Proc. Natl. Acad. Sci.* **110**, E1112–E1121 (2013).
49. Greene, S. B., Gunaratne, P. H., Hammond, S. M. & Rosen, J. M. A putative role for microRNA-205 in mammary epithelial cell progenitors. *J. Cell Sci.* **123**, 606–618 (2010).
50. Gandellini, P. *et al.* miR-205 regulates basement membrane deposition in human prostate: Implications for cancer development. *Cell Death Differ.* **19**, 1750–1760 (2012).
51. Wang, D. *et al.* MicroRNA-205 controls neonatal expansion of skin stem cells by modulating the PI (3) K pathway. *Nat. Cell Biol.* **15**, 1153–1163 (2013).
52. Yu, J. *et al.* MicroRNA-205 promotes keratinocyte migration via the lipid phosphatase SHIP2. *FASEB J.* **24**, 3950 (2010).
53. Nateghi, B., Emadi, F., Eghbali, M., Pezeshki, P. & Eshaghiyan, A. Circulating miR-193b-3p and miR-376a-3p involved in Iranian patients with multiple sclerosis. *Int. Biol. Biomed. J.* **5**, 24–28 (2019).
54. Wang, Y. *et al.* Circular RNA circDVL1 inhibits clear cell renal cell carcinoma progression through the miR-412-3p/PCDH7 axis. *Int. J. Biol. Sci.* **18**, 1491 (2022).
55. Lau, P., Frigerio, C. S. & De Strooper, B. Variance in the identification of microRNAs deregulated in Alzheimer's disease and possible role of lincRNAs in the pathology: The need of larger datasets. *Ageing Res. Rev.* **17**, 43–53 (2014).



56. Lim, Y. *et al.* Exploration of alcohol use disorder-associated brain miRNA–mRNA regulatory networks. *Transl. Psychiatry* **11**, 504 (2021).
57. Kumar, S. & Reddy, P. H. Are circulating microRNAs peripheral biomarkers for Alzheimer's disease?. *Biochim. Biophys. Acta (BBA)-Mol. Basis Dis.* **1862**, 1617–1627 (2016).
58. Skalsky, R. L. *et al.* The viral and cellular microRNA targetome in lymphoblastoid cell lines. *PLoS Pathog.* **8**, e1002484 (2012).
59. Cory, S. & Adams, J. M. The Bcl2 family: Regulators of the cellular life-or-death switch. *Nat. Rev. Cancer* **2**, 647–656 (2002).
60. Akhtar, R. S., Ness, J. M. & Roth, K. A. Bcl-2 family regulation of neuronal development and neurodegeneration. *Biochim. Biophys. Acta (BBA)-Mol. Cell Res.* **1644**, 189–203 (2004).
61. Mu, X., He, J., Anderson, D. W., Springer, J. E. & Trojanowski, J. Q. Altered expression of bcl-2 and bax mRNA in amyotrophic lateral sclerosis spinal cord motor neurons. *Ann. Neurol.* **40**, 379–386 (1996).
62. Troost, D., Aten, J., Morsink, F. & De Jong, J. Apoptosis in amyotrophic lateral sclerosis is not restricted to motor neurons. Bcl-2 expression is increased in unaffected post-central gyrus. *Neuropathol. Appl. Neurobiol.* **21**, 498–504 (1995).
63. Kostic, V., Jackson-Lewis, V., de Bilbao, F., Dubois-Dauphin, M. & Przedborski, S. Bcl-2: Prolonging life in a transgenic mouse model of familial amyotrophic lateral sclerosis. *Science* **277**, 559–563 (1997).
64. Azzouz, M. *et al.* Increased motoneuron survival and improved neuromuscular function in transgenic ALS mice after intraspinal injection of an adeno-associated virus encoding Bcl-2. *Hum. Mol. Genet.* **9**, 803–811 (2000).
65. Pasinelli, P. *et al.* Amyotrophic lateral sclerosis-associated SOD1 mutant proteins bind and aggregate with Bcl-2 in spinal cord mitochondria. *Neuron* **43**, 19–30 (2004).
66. Pedrini, S. *et al.* ALS-linked mutant SOD1 damages mitochondria by promoting conformational changes in Bcl-2. *Hum. Mol. Genet.* **19**, 2974–2986 (2010).
67. Lees, J., Shneidman, P., Skuntz, S., Carden, M. & Lazzarini, R. The structure and organization of the human heavy neurofilament subunit (NF-H) and the gene encoding it. *EMBO J.* **7**, 1947–1955 (1988).
68. Xie, H., Zhao, Y., Caramuta, S., Larsson, C. & Lui, W.-O. *miR-205 Expression Promotes Cell Proliferation and Migration of Human Cervical Cancer Cells.* (2012).
69. Sasaki, T. *et al.* In vivo and in vitro phosphorylation at Ser-493 in the glutamate (E)-segment of neurofilament-H subunit by glycogen synthase kinase 3 $\beta$ . *J. Biol. Chem.* **277**, 36032–36039 (2002).
70. Kim, M. S. *et al.* Neurofilament heavy polypeptide regulates the Akt- $\beta$ -catenin pathway in human esophageal squamous cell carcinoma. *PLoS ONE* **5**, e9003 (2010).
71. Lin, F. *et al.* Sequencing of neurofilament genes identified NEFH Ser787Arg as a novel risk variant of sporadic amyotrophic lateral sclerosis in Chinese subjects. *BMC Med. Genomics* **14**, 1–8 (2021).
72. Theunissen, F. *et al.* Intronic NEFH variant is associated with reduced risk for sporadic ALS and later age of disease onset. *Sci. Rep.-UK* **12**, 14739 (2022).
73. Campos-Melo, D., Hawley, Z. C. & Strong, M. J. Dysregulation of human NEFM and NEFH mRNA stability by ALS-linked miRNAs. *Mol. Brain* **11**, 1–9 (2018).
74. Pinacho, R. *et al.* Altered CSNK1E, FABP4 and NEFH protein levels in the dorsolateral prefrontal cortex in schizophrenia. *Schizophr. Res.* **177**, 88–97 (2016).
75. Iwamoto, K. *et al.* Decreased expression of NEFH and PCP4/PEP19 in the prefrontal cortex of alcoholics. *Neurosci. Res.* **49**, 379–385 (2004).
76. Yan, J. *et al.* A novel missense pathogenic variant in NEFH causing rare Charcot-Marie-Tooth neuropathy type 2CC. *Neurol. Sci.* **42**, 757–763 (2021).
77. Qiu, Y. *et al.* Emerging views of OPTN (optineurin) function in the autophagic process associated with disease. *Autophagy* **18**, 73–85 (2022).
78. Wild, P. *et al.* Phosphorylation of the autophagy receptor optineurin restricts Salmonella growth. *Science* **333**, 228–233 (2011).
79. Ying, H. & Yue, B. Y. Optineurin: The autophagy connection. *Exp. Eye Res.* **144**, 73–80 (2016).
80. Shen, X. *et al.* Processing of optineurin in neuronal cells. *J. Biol. Chem.* **286**, 3618–3629 (2011).
81. Yamano, K. *et al.* Critical role of mitochondrial ubiquitination and the OPTN–ATG9A axis in mitophagy. *J. Cell Biol.* **219**, 12144 (2020).
82. Sawa-Makarska, J. *et al.* Reconstitution of autophagosome nucleation defines Atg9 vesicles as seeds for membrane formation. *Science* **369**, eaaz7714 (2020).
83. Pensato, V. *et al.* Sorting rare ALS genetic variants by targeted re-sequencing panel in Italian patients: OPTN, VCP, and SQSTM1 variants account for 3% of rare genetic forms. *J. Clin. Med.* **9**, 412 (2020).
84. Canosa, A. *et al.* A familial amyotrophic lateral sclerosis pedigree discordant for a novel p. Glu46Asp heterozygous OPTN variant and the p. Ala5Val heterozygous SOD1 missense mutation. *J. Clin. Neurosci.* **75**, 223–225 (2020).
85. Mou, Y. *et al.* OPTN variants in ALS cases: A case report of a novel mutation and literature review. *Neurol. Sci.* **43**, 5391–5396 (2022).
86. Evans, C. S. & Holzbaur, E. L. Lysosomal degradation of depolarized mitochondria is rate-limiting in OPTN-dependent neuronal mitophagy. *Autophagy* **16**, 962–964 (2020).
87. Medchalmi, S., Tare, P., Sayyad, Z. & Swarup, G. A glaucoma-and ALS-associated mutant of OPTN induces neuronal cell death dependent on Tbk1 activity, autophagy and ER stress. *FEBS J.* **288**, 4576–4595 (2021).
88. Feldman, M., Benatar, M., Lavi, E. S., Latchman, K. & Singer, C. (AAN Enterprises, 2023).
89. Wang, C. *et al.* Isoforsythiaside attenuates Alzheimer's disease via regulating mitochondrial function through the PI3K/AKT pathway. *Int. J. Mol. Sci.* **21**, 5687 (2020).
90. Xiao, Q. *et al.* AMP-activated protein kinase-dependent autophagy mediated the protective effect of sonic hedgehog pathway on oxygen glucose deprivation-induced injury of cardiomyocytes. *Biochem. Biophys. Res. Commun.* **457**, 419–425 (2015).
91. Yin, X. *et al.* Downregulated AEG-1 together with inhibited PI3K/Akt pathway is associated with reduced viability of motor neurons in an ALS model. *Mol. Cell. Neurosci.* **68**, 303–313 (2015).

## Author contributions

W.Y.L. and H.P.L. initiated and supervised this study. J.C. and W.K.H. contributed to the acquisition, analysis, and interpretation of data. B.T.W. contributed to data analysis. All authors discussed and drafted the manuscript.

## Funding

This work was supported by grants from National Science and Technology Council of Taiwan (MOST 111-2320-B-039-032, MOST 111-2314-B-039-017-MY3) and grants from China Medical University & Hospital (CMU109-MF-110, CMU111-MF-50, CMU110-MF-92, DMR-112-125).

## Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-40879-y>.

**Correspondence** and requests for materials should be addressed to H.-P.L. or W.-Y.L.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023