



OPEN

## Dementia subtype prediction models constructed by penalized regression methods for multiclass classification using serum microRNA expression data

Yuya Asanomi<sup>1</sup>, Daichi Shigemizu<sup>1,2,3✉</sup>, Shintaro Akiyama<sup>1</sup>, Takashi Sakurai<sup>4,5</sup>, Kouichi Ozaki<sup>1,3</sup>, Takahiro Ochiya<sup>6,7</sup> & Shumpei Niida<sup>1✉</sup>

There are many subtypes of dementia, and identification of diagnostic biomarkers that are minimally-invasive, low-cost, and efficient is desired. Circulating microRNAs (miRNAs) have recently gained attention as easily accessible and non-invasive biomarkers. We conducted a comprehensive miRNA expression analysis of serum samples from 1348 Japanese dementia patients, composed of four subtypes—Alzheimer’s disease (AD), vascular dementia, dementia with Lewy bodies (DLB), and normal pressure hydrocephalus—and 246 control subjects. We used this data to construct dementia subtype prediction models based on penalized regression models with the multiclass classification. We constructed a final prediction model using 46 miRNAs, which classified dementia patients from an independent validation set into four subtypes of dementia. Network analysis of miRNA target genes revealed important hub genes, *SRC* and *CHD3*, associated with the AD pathogenesis. Moreover, *MCU* and *CASP3*, which are known to be associated with DLB pathogenesis, were identified from our DLB-specific target genes. Our study demonstrates the potential of blood-based biomarkers for use in dementia-subtype prediction models. We believe that further investigation using larger sample sizes will contribute to the accurate classification of subtypes of dementia.

### Abbreviations

AD	Alzheimer’s disease
CN	Cognitively normal elder control
CSF	Cerebrospinal fluid
DLB	Dementia with Lewy bodies
GSEA	Gene set enrichment analyses
miRNA	MicroRNA
NPH	Normal pressure hydrocephalus
VaD	Vascular dementia

Dementia is a syndrome characterized by loss of cognitive function, in particular memory, thinking, and behavioral abilities. The number of people worldwide with dementia among the elderly population is rapidly increasing, and is expected to reach 74.7 million in 2030 and 131.5 million in 2050<sup>1</sup>. The most common subtype of dementia

<sup>1</sup>Medical Genome Center, Research Institute, National Center for Geriatrics and Gerontology, 7-430 Morioka-cho, Obu, Aichi 474-8511, Japan. <sup>2</sup>Department of Medical Science Mathematics, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan. <sup>3</sup>RIKEN Center for Integrative Medical Sciences, Yokohama, Japan. <sup>4</sup>Center for Comprehensive Care and Research on Memory Disorders, National Center for Geriatrics and Gerontology, Obu, Aichi, Japan. <sup>5</sup>Department of Cognition and Behavior Science, Nagoya University Graduate School of Medicine, Nagoya, Aichi, Japan. <sup>6</sup>Division of Molecular and Cellular Medicine, Fundamental Innovative Oncology Core Center, National Cancer Center Research Institute, Tokyo, Japan. <sup>7</sup>Department of Molecular and Cellular Medicine, Institute of Medical Science, Tokyo Medical University, Tokyo, Japan. ✉email: daichi@ncgg.go.jp; sniida@ncgg.go.jp

is Alzheimer's disease (AD), followed by vascular dementia (VaD), and dementia with Lewy bodies (DLB). The diagnoses of these subtypes of dementia are based on clinical criteria for each subtype<sup>2–6</sup>. However, it is difficult to differentiate between dementia subtypes, because many clinical features overlap. For example, patients with normal pressure hydrocephalus (NPH) present with similar symptoms to AD, although NPH is often curable by shunt surgery to remove accumulated cerebrospinal fluid (CSF)<sup>7</sup>. Therefore, the identification of new biomarkers for more efficient diagnosis is urgently required.

Imaging-based techniques, including positron emission tomography scans for detection of amyloid beta (A $\beta$ ) deposition or tau tracers, and volumetric magnetic resonance imaging for determination of hippocampal or medial temporal lobe atrophy, are currently used for diagnosis of dementia subtypes; however, these techniques are not suitable for initial screening due to their high cost<sup>8</sup>. In addition, while several CSF biomarkers, such as A $\beta$ <sub>1–42</sub>, total tau (T-tau), and phosphorylated tau 181 (P-tau<sub>181</sub>), are effective for characterizing AD, these are not suitable for initial screening due to their high invasiveness. Therefore, blood-based biomarkers are the most attractive option as candidate biomarkers for practical clinical use due to their minimal invasiveness and cost effectiveness.

MicroRNAs (miRNAs) are small non-coding RNAs of 20–25 nucleotides in length that regulate gene expression by binding to complementary regions of messenger RNAs<sup>9</sup>. Many circulating miRNAs in blood have been identified as promising biomarkers for cancer diagnosis<sup>10–15</sup> and dementia diagnosis<sup>16–22</sup>. However, the prediction models constructed for dementia using miRNAs are specific to each subtype<sup>16–22</sup> and the development of a multiclass classification prediction model of dementia subtypes is required.

Here, we conducted a comprehensive serum miRNA expression analysis of 1348 Japanese dementia patients composed of four subtypes (AD, VaD, DLB, and NPH) and 246 control subjects. Using this data, we constructed various dementia subtype prediction models based on penalized regression models with multiclass classification. Our final prediction model classified dementia patients into four subtypes of dementia using 46 miRNAs. Furthermore, network-based meta-analysis of the miRNA target genes revealed several important hub genes associated with the pathogenesis of dementia subtypes. We believe that further investigation using a larger sample size will contribute to accurate classification of dementia subtypes.

## Methods

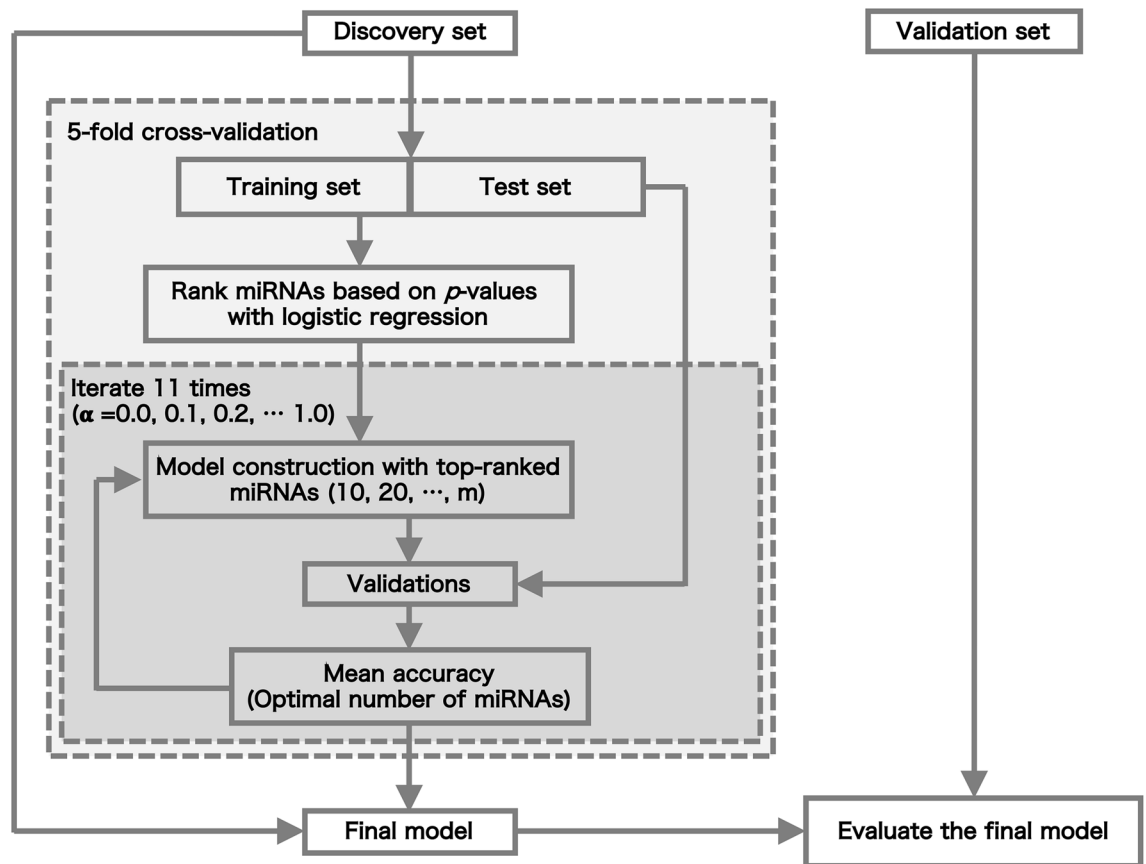
**Clinical samples.** All 1594 serum samples and clinical information, including sex, age, and *APOE*  $\epsilon$ 4 genotypes, were obtained from the National Center for Geriatrics and Gerontology (NCGG) Biobank. Of these samples, 1009 were from patients with AD, 89 were from patients with VaD, 166 were from patients with DLB, and 84 were from patients with NPH; 246 samples were from cognitively normal elder controls (CN). The AD patients were diagnosed based on the criteria developed by the National Institute on Aging and the Alzheimer's Association (NIA-AA)<sup>2,3</sup>; the VaD patients were diagnosed using the criteria in the report of the NINDS-AIREN International Workshop<sup>4</sup>; and the DLB and NPH patients were diagnosed based on the criteria in the fourth report of the DLB Consortium<sup>5</sup> and the guidelines for the management of NPH<sup>6</sup>, respectively. The CN samples had subjective cognitive complaints, but normal cognition on the neuropsychological assessment with a comprehensive neuropsychological test (Mini-Mental State Examination score, > 25). All individuals were 60 years of age or older. This study was approved by the ethics committee of the NCGG. The design and performance of the study were clearly described in a research protocol. Participation was voluntary and all participants completed informed consent in writing before registering to NCGG Biobank. All methods in this study were performed in accordance with the Declaration of Helsinki.

**miRNA-microarray assay.** Total RNAs were extracted from 300  $\mu$ L serum samples by using a 3D-Gene RNA extraction reagent (Toray Industries, Inc., Kanagawa, Japan). Comprehensive miRNA expression analysis was performed using the 3D-Gene system and a 3D-Gene Human miRNA Oligo Chip (Toray Industries, Inc.), which was designed to detect 2565 of the miRNA sequences registered in miRBase (release 21, <http://www.mirbase.org/>). All microarray data used in this study are publicly available through the Gene Expression Omnibus database accession numbers: GSE120584 and GSE167559 (<https://www.ncbi.nlm.nih.gov/geo/>). The miRNA expression normalization was described in our previous study<sup>20</sup>.

**Construction and evaluation of the dementia-type risk prediction model.** All data were strictly separated into a discovery set and validation set (Fig. 1). Four-fifths of the discovery set was used to calculate *p*-values in each cross-validation step. The *p*-values corresponding to the miRNAs were calculated with a logistic regression model between dementia and CN with adjustments for three covariates: age, sex, and the number of *APOE*  $\epsilon$ 4 alleles. The top-ranked *m* miRNAs were selected stepwise (*m* = 10, 20, ...). Using a combination of the *m* miRNAs and three clinical variables, risk prediction models were constructed based on penalized regression methods: ridge regression, elastic net, and lasso. Let  $X_i = (X_{i,1}, \dots, X_{i,p})$  be the values of the pre-selected top-ranked *m* miRNAs for a subject *i* and let  $l(\beta; \gamma_i, X_i)$  be the logistic log-likelihood:

$$l(\beta; \gamma_i, X_i) - \lambda P_\alpha(\beta),$$

where  $P_\alpha(\beta) = (1 - \alpha)\frac{1}{2}\beta^2 + \alpha|\beta|$  and  $\alpha$  is set to 1 (lasso), 0 (ridge regression), or 0.1 to 0.9 at 0.1 intervals (elastic net), and the optimal penalty parameter  $\lambda$  is selected using fivefold cross-validation. This process was repeated 5 times (fivefold cross-validation). Based on the average accuracy, we determined the optimal number of miRNAs and  $\alpha$  for model construction. The final model was constructed using the entire discovery set and the adjusted model was evaluated on the independent validation set. The multiclass classification model used in this study was implemented using the *glmnet* package (version 2.0–18)<sup>23</sup> with the function family set to "multinomial" in the statistical software R<sup>24</sup>.



**Figure 1.** Outline of construction of dementia subtype prediction model and validation.

**Network analysis for the target genes of miRNAs.** Up- and down-regulated miRNAs were defined based on the coefficients of the multiclass dementia subtype prediction model. The miRNAs with positive coefficients were assigned to up-regulated miRNAs, and those with negative coefficients were assigned to down-regulated miRNAs. The miRNA target gene annotation was conducted using the microRNA Target Prediction and Functional Study Database (miRDB)<sup>25</sup>. Only target genes with a prediction score (assigned by MirTarget V3 based on miRBase 21) of greater than 90 were used in this study. Based on the target genes, protein–protein interaction (PPI) network analyses were implemented using the web-based tool NetworkAnalyst 3.0 (<https://www.networkanalyst.ca/>)<sup>26</sup> together with the International Molecular Exchange Consortium Interactome database<sup>27</sup>; we set the betweenness cutoff, which indicates the number of shortest paths through a node, to 25 and used the minimum network filter. The PPI networks were visualized using Cytoscape (version 3.7.1)<sup>28</sup>. We identified nodes with degree > 100 as hub nodes. We identified statistically significant KEGG pathways and biological process gene ontologies (GOs) by using NetworkAnalyst. The *P*-value was corrected for multiple testing using False Discovery Rate (FDR), and the statistical significance threshold was set at FDR < 0.05.

**Ethics approval and consent to participate.** This study was approved by the ethics committee of the NCGG. The design and performance of the current study, which involves human subjects, were clearly described in a research protocol. All participation was voluntary, and participants completed informed consent in writing before registering with the NCGG Biobank.

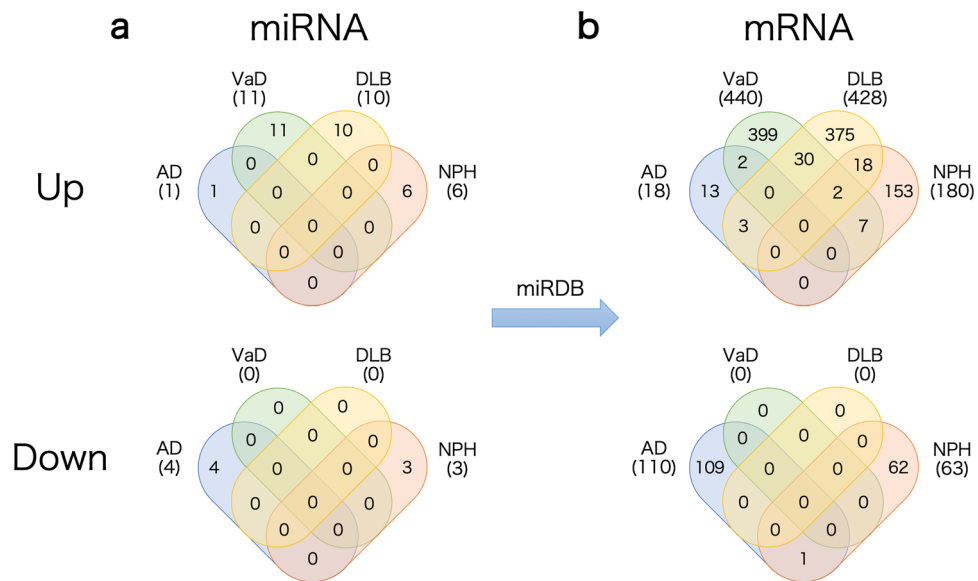
## Results

**Subjects.** A total of 1594 Japanese people comprising 1009 AD, 89 VaD, 166 DLB, 84 NPH, and 246 CN subjects, were enrolled in this study in the NCGG Biobank. The male:female ratio of all individuals was 1:1.68 with an average age of 78.0 years (SD, 6.8). The allele frequency of the C allele of rs429358, which defines the *APOE*  $\epsilon 4$  phenotype, was 0.203 (Table 1). We also split the participants into a discovery set and a validation set. This separation was stratified on age and sex such that there was a similar distribution of these features in the discovery and validation sets.

**Dementia subtype prediction models.** We constructed dementia subtype prediction models based on penalized regression methods for multiclass classification (Fig. 1). We carefully examined the sample size of each dementia subtype in the discovery data set, as the imbalance of sample size in multiclass classification is a well-known issue in the field of machine learning. The detailed procedures of four models, Model A–D, are described in Supplementary Note and Supplementary information (Figs. S1, S2, and Tables S1–S5).

Group	No. of samples	Male:Female	Age (mean $\pm$ 1 S.D.)	rs429358 <sup>a</sup>			
				TT	TC	CC	MAF <sup>b</sup>
AD	1009	1:2.33	79.3 $\pm$ 6.2	565	376	68	0.254
VaD	89	1:0.68	79.1 $\pm$ 6.4	67	21	1	0.129
DLB	166	1:1.52	79.5 $\pm$ 6.0	115	48	3	0.163
NPH	84	1:0.95	78.9 $\pm$ 6.1	69	14	1	0.095
CN	246	1:0.91	71.1 $\pm$ 6.2	207	36	3	0.085
Total	1594	1:1.68	78.0 $\pm$ 6.8	1023	495	76	0.203

**Table 1.** Sample characteristics of each dementia subtype. <sup>a</sup>The C allele of rs429358 defines the *APOE*  $\epsilon$ 4 phenotype, while the T allele defines wild-type,  $\epsilon$ 3. <sup>b</sup>MAF Minor allele frequency.

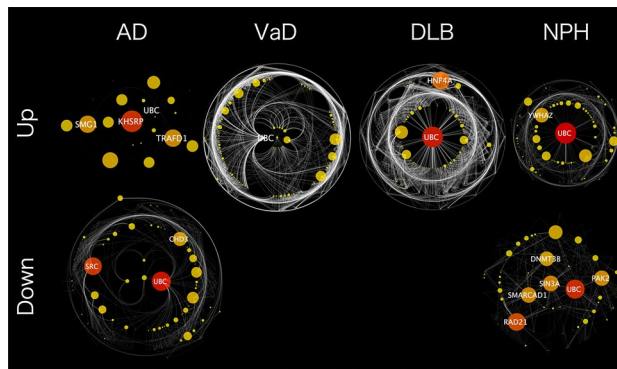


**Figure 2.** Effective miRNAs and genes used in our dementia-type prediction model. (a) The number of up- and down-regulated miRNAs used for the final models for each dementia subtype. (b) Target genes for each miRNA used for each model; the target genes were predicted by using annotation in the miRDB data.

To eliminate the bias of sample size among dementia subtypes, we constructed a final model using the same number of samples in each dementia subtype. We selected 50 samples, composed of 25 males and 25 females, from four dementia subtypes and CN. In total, 250 samples were used for a discovery data set and the remaining samples were used for a validation data set (Supplementary Table S3, Dataset 3). Through the fivefold cross-validation, the maximum mean accuracy for dementia subtypes was achieved with a parameter combination of  $(m, \alpha) = (200, 0.9)$  (Supplementary Fig. S3). The final prediction model, Model D (Supplementary Table S6), was constructed with the optimal parameters detected using the entire discovery set and achieved an accuracy of 0.398 in the validation data set and a mean accuracy of 0.374 for dementia subtypes (Supplementary Table S3) when using 46 miRNAs.

**PPI network analyses using miRNA target genes.** According to the annotation in miRDB, the 46 miRNAs used in Model D were predicted to target 1303 genes. Of these 1303 genes, the numbers of genes related to up- and down-regulated miRNAs were, respectively, 18 and 110 in AD, 440 and 0 in VaD, 428 and 0 in DLB, and 180 and 63 in NPH (Fig. 2). To elucidate functional modules from the target genes, PPI network analyses were performed using NetworkAnalyst for each dementia subtype; PPI network visualization was conducted with Cytoscape software (Fig. 3). Several interesting hub genes were identified in each dementia subtype. Of them, *SRC* and *CHD3*, which are targets of down-regulated miRNAs in AD, are particularly interesting. The *SRC* gene is reported to be associated with AD<sup>29–31</sup>. The *CHD3* gene is reported to be involved in transcription of the *PSEN1* gene, the most common cause of familial Alzheimer's disease<sup>32–34</sup>. While the *UBC* gene was observed in all dementia subtypes networks, it was not one of the 1303 target genes that we detected.

We further performed gene set enrichment analyses (GSEA) using NetworkAnalyst for each dementia subtype. A statistically significant KEGG pathway (Ras signaling pathway, hsa04014, FDR = 0.0083 in VaD) and four statistically significant GO terms (Nucleobase-containing compound transport, GO:0015931, FDR = 0.029



**Figure 3.** PPI network analysis for the genes that were targeted by the miRNAs. The gene symbol is displayed for nodes with degree > 100.

Gene set	Group	Up-/Down -regulated miRNAs	Pathway/GO term	No. of genes (Hits/Total)	<i>p</i> -value	FDR <sup>a</sup>
KEGG	VaD	Up	Ras signaling pathway (KEGG: hsa04014)	17/232	$2.6 \times 10^{-5}$	0.0083
GO	AD	Up	Nucleobase-containing compound transport (GO:0015931)	4/175	$3.5 \times 10^{-5}$	0.029
		Up	RNA export from nucleus (GO:0006405)	3/86	$1.1 \times 10^{-4}$	0.046
	DLB	Up	Homophilic cell adhesion (GO:0007156)	17/139	$6.5 \times 10^{-8}$	$5.3 \times 10^{-5}$
		Up	Cell–cell adhesion (GO:0098609)	30/461	$1.6 \times 10^{-6}$	$6.6 \times 10^{-4}$

**Table 2.** Gene set enrichment analyses using the target genes of the miRNAs. <sup>a</sup>FDR false discovery rate.

in AD; RNA export from nucleus, GO:0006405, FDR = 0.046 in AD; Homophilic cell adhesion, GO:0007156, FDR =  $5.3 \times 10^{-5}$  in DLB; Cell–cell adhesion, GO:0098609, FDR =  $6.6 \times 10^{-4}$  in DLB) were identified in the target genes of up-regulated miRNAs (Table 2).

## Discussion

We report a prediction model for multiclass classification of dementia subtypes using serum miRNAs. Most previous research has focused on developing prediction models for specific dementia subtypes<sup>16–22</sup>. Shigemizu et al. reported separate effective blood miRNA-based risk prediction models for several dementia subtypes<sup>20,21</sup>; however, such individual dementia subtype prediction models would likely predict multiple dementia subtypes for each patient. To address this issue, we propose a multiclass classification dementia subtype prediction model. Although our models could not provide sufficient accuracy, a further, larger cohort study is expected to improve the prediction accuracy. In particular, an improvement of the prediction accuracy can be expected in patients with VaD and NPH as the number of affected patients in this study was small.

Network analyses were performed using the target genes of the 46 miRNAs involved in our model that used equal sample sizes for the four subtypes, Model D. Two interesting hub genes, *SRC* and *CHD3*, were detected from the PPI network of the target genes of down-regulated miRNAs in AD. Accumulation of A $\beta$  in the brain is strongly associated with the onset of AD. The constitutively active form of Src increases A $\beta$  generation, and the inhibition of the kinase activity of Src reduces the generation of A $\beta$  in cells stably expressing *APP*<sup>29</sup>. Src also regulates the phosphorylation of Mint, which has an important role in the production and the secretion of A $\beta$ <sup>30,31</sup>. *CHD3* suppresses transcription of the familial AD causative gene *PSEN1*<sup>32,33</sup> and may potentially rescue Ca<sup>2+</sup>-signaling defects in familial AD patients and prevent neuronal apoptosis<sup>34</sup>. While *UBC* appeared in all networks, there were no miRNAs targeting *UBC* in our model. This may be because the housekeeping gene *UBC* (Ubiquitin C) has a key role in various pathways, such as DNA repair, cell cycle regulation, kinase modification, endocytosis, and regulation of other cell signaling pathways.

Dementia-subtype-specific target genes could be effective in distinguishing between the four dementia subtypes. Two genes, *MCU* and *CASP3*, that are associated with DLB pathogenesis<sup>35–37</sup>, were identified from the target genes of up-regulated miRNAs in DLB. Calcium uptake via the mitochondrial calcium uniporter (MCU) complex and mitochondrial calcium overload play a key role in neurodegenerative diseases, including DLB<sup>35,36</sup>. Neuronal accumulation of  $\alpha$ -synuclein is a major feature of DLB. Desplats et al. reported that neuron-to-neuron transmission of  $\alpha$ -synuclein triggers activation of caspase 3, which is coded by *CASP3*, and causes neuronal cell death<sup>37</sup>.

The Ras signaling pathway (hsa04014) was identified through GSEA using target genes of up-regulated miRNAs in VaD. The Ras signaling pathway was also identified from target genes of down-regulated miRNAs in AD, but the result was not statistically significant after FDR correction. Kirouac et al. reported that the expression of Ras is associated with A $\beta$  production in human AD brains<sup>38</sup>. Therefore, further study of genes involved in the Ras signaling pathway could contribute to understanding the difference between AD and VaD. In addition, four statistically significant GO terms (Nucleobase-containing compound transport, RNA export from nucleus, Homophilic cell adhesion, and Cell–cell adhesion) were identified from target genes of up-regulated miRNAs in AD and DLB. The association between these terms and dementia is currently unknown. These GO terms, however, may provide important clues to the underlying mechanisms for AD and DLB pathogenesis.

Blood-based biomarkers, which form the basis of our prediction models are attractive because they are minimally invasive, cost-effective, and easy to reproduce; however, the strategy we present here has a limitation. It is not applicable to mixed dementia cases, in which patients exhibit more than one subtype of dementia simultaneously<sup>39,40</sup>. As long as this issue is unresolved, it may be difficult to apply our findings to practical clinical use.

## Conclusions

We developed dementia subtype prediction models by using penalized regression methods for multiclass classification and serum microRNA expression data. Network analysis of the miRNA target genes revealed several important hub genes associated with AD pathogenesis. Moreover, we found several genes associated with DLB pathogenesis, from our DLB-specific target genes. Our study demonstrates the potential of blood-based biomarkers for dementia subtype prediction models. We believe that further investigation using a larger sample size will contribute to a more accurate classification of dementia subtypes.

## Data availability

All microarray data used in this study are publicly available through the Gene Expression Omnibus database accession number GSE120584 and GSE167559 (<https://www.ncbi.nlm.nih.gov/geo/>).

Received: 26 March 2021; Accepted: 12 October 2021

Published online: 22 October 2021

## References

- Prince, M. et al. *World Alzheimer Report 2015: The Global Impact of Dementia: An Analysis of Prevalence, Incidence, Cost And trends* (Published by Alzheimer's Disease International, 2015).
- McKhann, G. M. et al. The diagnosis of dementia due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* **7**, 263–269 (2011).
- Albert, M. S. et al. The diagnosis of mild cognitive impairment due to Alzheimer's disease: Recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's Dement.* **7**, 270–279 (2011).
- Román, G. C. et al. Vascular dementia: Diagnostic criteria for research studies. Report of the NINDS-AIREN International Workshop. *Neurology* **43**, 250–260 (1993).
- McKeith, I. G. et al. Diagnosis and management of dementia with Lewy bodies. *Neurology* **89**, 88–100 (2017).
- Mori, E. et al. Guidelines for management of idiopathic normal pressure hydrocephalus. *Neurol. Med. Chir.* **52**, 775–809 (2012).
- Adams, R. D., Fisher, C. M., Hakim, S., Ojemann, R. G. & Sweet, W. H. Symptomatic occult hydrocephalus with normal cerebrospinal-fluid pressure: A treatable syndrome. *N. Engl. J. Med.* **273**, 117–126 (1965).
- Jack, C. R. et al. NIA-AA research framework: Toward a biological definition of Alzheimer's disease. *Alzheimers. Dement.* **14**, 535–562 (2018).
- Kim, V. N., Han, J. & Siomi, M. C. Biogenesis of small RNAs in animals. *Nat. Rev. Mol. Cell Biol.* **10**, 126–139 (2009).
- Kosaka, N., Iguchi, H. & Ochiya, T. Circulating microRNA in body fluid: A new potential biomarker for cancer diagnosis and prognosis. *Cancer Sci.* **101**, 2087–2092 (2010).
- Shimomura, A. et al. Novel combination of serum microRNA for detecting breast cancer in the early stage. *Cancer Sci.* **107**, 326–334 (2016).
- Yokoi, A. et al. Integrated extracellular microRNA profiling for ovarian cancer screening. *Nat. Commun.* **9**, 4319 (2018).
- Matsuzaki, J. & Ochiya, T. Circulating microRNAs and extracellular vesicles as potential cancer biomarkers: A systematic review. *Int. J. Clin. Oncol.* **22**, 413–420 (2017).
- Falzone, L. et al. Identification of novel microRNAs and their diagnostic and prognostic significance in oral cancer. *Cancers* **11**, 610 (2019).
- Jiang, X. et al. Serum microRNA expression signatures identified from genome-wide microRNA profiling serve as novel noninvasive biomarkers for diagnosis and recurrence of bladder cancer. *Int. J. Cancer* **136**, 854–862 (2015).
- Cogswell, J. P. et al. Identification of miRNA changes in Alzheimer's disease brain and CSF yields putative biomarkers and insights into disease pathways. *J. Alzheimers. Dis.* **14**, 27–41 (2008).
- Sørensen, S. S., Nygaard, A.-B. & Christensen, T. miRNA expression profiles in cerebrospinal fluid and blood of patients with Alzheimer's disease and other types of dementia: An exploratory study. *Transl. Neurodegener.* **5**, 6 (2016).
- Angelucci, F. et al. MicroRNAs in Alzheimer's disease: Diagnostic markers or therapeutic agents? *Front. Pharmacol.* **10**, 665 (2019).
- Kayano, M. et al. Plasma microRNA biomarker detection for mild cognitive impairment using differential correlation analysis. *Biomark. Res.* **4**, 22 (2016).
- Shigemizu, D. et al. Risk prediction models for dementia constructed by supervised principal component analysis using miRNA expression data. *Commun. Biol.* **2**, 77 (2019).
- Shigemizu, D. et al. A comparison of machine learning classifiers for dementia with Lewy bodies using miRNA expression data. *BMC Med. Genomics* **12**, 150 (2019).
- Ragusa, M. et al. miRNAs plasma profiles in vascular dementia: Biomolecular data and biomedical implications. *Front. Cell. Neurosci.* **10**, 51 (2016).
- Friedman, J., Hastie, T. & Tibshirani, R. Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* **33**, 1–22 (2010).

24. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical (Computing, Vienna, Austria). (2019).
25. Liu, W. & Wang, X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biol.* **20**, 18 (2019).
26. Zhou, G. *et al.* NetworkAnalyst 3.0: A visual analytics platform for comprehensive gene expression profiling and meta-analysis. *Nucleic Acids Res.* **47**, W234–W241 (2019).
27. Breuer, K. *et al.* InnateDB: Systems biology of innate immunity and beyond: Recent updates and continuing curation. *Nucleic Acids Res.* **41**, D1228–D1233 (2013).
28. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
29. Gianni, D. *et al.* Platelet-derived growth factor induces the  $\beta$ - $\gamma$ -secretase-mediated cleavage of Alzheimer's amyloid precursor protein through a Src-Rac-dependent pathway. *J. Biol. Chem.* **278**, 9290–9297 (2003).
30. Chaufty, J., Sullivan, S. E. & Ho, A. Intracellular amyloid precursor protein sorting and amyloid- $\beta$  secretion are regulated by Src-mediated phosphorylation of Mint2. *J. Neurosci.* **32**, 9613–9625 (2012).
31. Dunning, C. J. R. *et al.* Multisite tyrosine phosphorylation of the N-terminus of Mint1/X11a by Src kinase regulates the trafficking of amyloid precursor protein. *J. Neurochem.* **137**, 518–527 (2016).
32. Pastorcic, M. & Das, H. K. The C-terminal region of CHD3/ZFH interacts with the CIDD region of the Ets transcription factor ERM and represses transcription of the human presenilin 1 gene. *FEBS J.* **274**, 1434–1448 (2007).
33. Lee, S. & Das, H. K. Transcriptional regulation of the presenilin-1 gene controls gamma-secretase activity. *Front. Biosci. (Elite Ed)* **2**, 22–35 (2010).
34. Das, H. K., Tchedre, K. & Mueller, B. Repression of transcription of presenilin-1 inhibits  $\gamma$ -secretase independent ER Ca<sup>2+</sup> leak that is impaired by FAD mutations. *J. Neurochem.* **122**, 487–500 (2012).
35. Verma, M. *et al.* Mitochondrial calcium dysregulation contributes to dendrite degeneration mediated by PD/LBD-associated LRRK2 mutants. *J. Neurosci.* **37**, 11151–11165 (2017).
36. Verma, M., Wills, Z. & Chu, C. T. Excitatory dendritic mitochondrial calcium toxicity: Implications for Parkinson's and other neurodegenerative diseases. *Front. Neurosci.* **12**, 523 (2018).
37. Desplats, P. *et al.* Inclusion formation and neuronal cell death through neuron-to-neuron transmission of  $\alpha$ -synuclein. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 13010–13015 (2009).
38. Kirouac, L., Rajic, A. J., Cribbs, D. H. & Padmanabhan, J. Activation of Ras-ERK signaling and GSK-3 by amyloid precursor protein and amyloid beta facilitates neurodegeneration in Alzheimer's disease. *eneuro* <https://doi.org/10.1523/ENEURO.0149-16.2017> (2017).
39. Zekry, D., Hauw, J.-J. & Gold, G. Mixed dementia: Epidemiology, diagnosis, and treatment. *J. Am. Geriatr. Soc.* **50**, 1431–1438 (2002).
40. Jellinger, K. A. & Attems, J. Neuropathological evaluation of mixed dementia. *J. Neurol. Sci.* **257**, 80–87 (2007).

## Acknowledgements

We thank the staff members of NCGG Medical Genome Center for their contribution to providing the serum samples and clinical information. All miRNA-microarray data were provided by New Frontiers Research Laboratories, Toray Industries, Inc., Kanagawa, Japan.

## Author contributions

Y.A. develop the models and performed the analyses; D.S., S.A., and S.N. provided the technical assistance; T.S., T.O., and S.N. contributed to data acquisition; Y.A. and D.S. wrote the manuscript; D.S., T.O., K.O., and S.N. organized this work. All authors read and approved the final manuscript.

## Funding

This study was supported by the Japan Foundation for Aging and Health and the Takeda Science Foundation (to D.S.); the Research Funding for Longevity Sciences from the National Center for Geriatrics and Gerontology (29-45 to K.O. and 30-29 to D.S., and 20-10/21-22 to S.N.); a grant for Research on Dementia from the Japanese Ministry of Health, Labor and Welfare (to K.O.); a grant from the Japan Agency for Medical Research and Development (AMED) (Grant Number JP18kk0205009 to S.N.); and a grant for Development of Diagnostic Technology for Detection of miRNA in Body Fluids from the Japan Agency for Medical Research and Development and New Energy and Industrial Technology Development Organization (Grant Number JP17ae0101013 to S.N.).

## 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/s41598-021-00424-1>.

**Correspondence** and requests for materials should be addressed to D.S. or S.N.

**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) 2021