



**HAL**  
open science

## MicroRNA signatures in genetic frontotemporal dementia and amyotrophic lateral sclerosis

Virgilio Kmetzsch, Morwena Latouche, Dario Saracino, Daisy Rinaldi, Agnès Camuzat, Thomas Gareau, Isabelle Le Ber, Olivier Colliot, Emmanuelle Becker

► **To cite this version:**

Virgilio Kmetzsch, Morwena Latouche, Dario Saracino, Daisy Rinaldi, Agnès Camuzat, et al.. MicroRNA signatures in genetic frontotemporal dementia and amyotrophic lateral sclerosis. *Annals of Clinical and Translational Neurology*, 2022, pp.1-14. 10.1002/acn3.51674 . hal-03826747

**HAL Id: hal-03826747**

**<https://hal.science/hal-03826747>**

Submitted on 24 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.


L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

## RESEARCH ARTICLE

# MicroRNA signatures in genetic frontotemporal dementia and amyotrophic lateral sclerosis

Virgilio Kmetzsch<sup>1,2</sup> , Morwena Latouche<sup>3</sup>, Dario Saracino<sup>1,4,5</sup>, Daisy Rinaldi<sup>3,4,5</sup>, Agnès Camuzat<sup>3,6</sup>, Thomas Gareau<sup>3</sup>, the French Research Network on FTD/ALS, Isabelle Le Ber<sup>3,4,5,7</sup>, Olivier Colliot<sup>1</sup> & Emmanuelle Becker<sup>2</sup>

<sup>1</sup>Sorbonne Université, Institut du Cerveau – Paris Brain Institute – ICM, CNRS, Inria, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013, Paris, France

<sup>2</sup>Univ Rennes, Inria, CNRS, IRISA, F-35000, Rennes, France

<sup>3</sup>Sorbonne Université, Institut du Cerveau – Paris Brain Institute – ICM, CNRS, Inserm, AP-HP, Hôpital de la Pitié Salpêtrière, F-75013, Paris, France

<sup>4</sup>Centre de référence des démences rares ou précoces, IM2A, Département de Neurologie, AP-HP - Hôpital Pitié-Salpêtrière, Paris, France

<sup>5</sup>Département de Neurologie, AP-HP - Hôpital Pitié-Salpêtrière, Paris, France

<sup>6</sup>EPHE, PSL Research University, Paris, France

<sup>7</sup>Paris Brain Institute – Institut du Cerveau – ICM, FrontLab, Paris, France

## Correspondence

Emmanuelle Becker, Université de Rennes 1 – Campus Beaulieu, 263 Avenue Général Leclerc, 35042 Rennes, France. Tel: 0033 2 9984 2567; Fax: +33 2 99 84 71 71; E-mail: emmanuelle.becker@univ-rennes1.fr

## Funding Information

The research leading to these results has received funding from the program “Investissements d’avenir” ANR-10-IAIHU-06, from the French government under management of Agence Nationale de la Recherche as part of the “Investissements d’avenir” program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute), from Agence Nationale de la Recherche/DGOS (project ANR-PRTS PREV-DEMALS, grant number ANR-14-CE15-0016-07, promotion Assistance Publique – Hôpitaux de Paris), from the Programme Hospitalier de Recherche Clinique (PHRC) Predict-PGRN (to ILB, promotion by Assistance Publique – Hôpitaux de Paris), from Fondation Plan Alzheimer (to ILB), from the Inria Project Lab Program (project Neuromarkers), from Fondation Vaincre Alzheimer FR-17035 and from the Institut Français de Bioinformatique (ANR-11-INSB-0013).

Received: 15 July 2022; Revised: 12 September 2022; Accepted: 20 September 2022

doi: 10.1002/acn3.51674

## Introduction

Frontotemporal dementia (FTD) is a neurodegenerative disease characterized by brain atrophy in the frontal and

## Abstract

**Objective:** MicroRNAs are promising biomarkers of frontotemporal dementia (FTD) and amyotrophic lateral sclerosis (ALS), but discrepant results between studies have so far hampered their use in clinical trials. We aim to assess all previously identified circulating microRNA signatures as potential biomarkers of genetic FTD and/or ALS, using homogeneous, independent validation cohorts of *C9orf72* and *GRN* mutation carriers. **Methods:** 104 individuals carrying a *C9orf72* or a *GRN* mutation, along with 31 controls, were recruited through the French research network on FTD/ALS. All subjects underwent blood sampling, from which circulating microRNAs were extracted. We measured differences in the expression levels of 65 microRNAs, selected from 15 published studies about FTD or ALS, between 31 controls, 17 *C9orf72* presymptomatic subjects, and 29 *C9orf72* patients. We also assessed differences in the expression levels of 30 microRNAs, selected from five studies about FTD, between 31 controls, 30 *GRN* presymptomatic subjects, and 28 *GRN* patients. **Results:** More than half (35/65) of the selected microRNAs were differentially expressed in the *C9orf72* cohort, while only a small proportion (5/30) of microRNAs were differentially expressed in the *GRN* cohort. In multivariate analyses, only individuals in the *C9orf72* cohort could be adequately classified (ROC AUC up to 0.98 for controls versus presymptomatic subjects, 0.94 for controls versus patients, and 0.77 for presymptomatic subjects versus patients) with some of the signatures. **Interpretation:** Our results suggest that previously identified microRNAs using sporadic or mixed cohorts of FTD and ALS patients could potentially serve as biomarkers of *C9orf72*-associated disease, but not *GRN*-associated disease.

temporal lobes, causing severe changes in personality and social behavior.<sup>1</sup> The most prevalent genetic causes of FTD are GGGGCC repeat expansions in the *C9orf72* gene and mutations in the *GRN* gene.<sup>2,3</sup> FTD shares disease

pathways with amyotrophic lateral sclerosis (ALS), a debilitating motor neuron disease that causes progressive motor deficit and muscle wasting.<sup>4</sup> The *C9orf72* hexanucleotide repeat expansion has been identified as the most common genetic cause of both familial FTD and ALS, as well as of their sporadic counterparts.<sup>2</sup>

There are currently no disease-modifying treatments that can stop the course of FTD or ALS. New therapeutic trials depend on robust progression biomarkers to assess treatment outcomes. The study of FTD/ALS genetic forms is particularly important, since asymptomatic mutation carriers may provide insights about the early disease stages, before any irreversible neuronal damage.<sup>5</sup>

Among the potential noninvasive biomarkers of neurodegenerative diseases, circulating microRNAs (miRNAs) constitute a promising approach.<sup>6</sup> MiRNAs are short non-coding RNAs that negatively regulate gene expression.<sup>7</sup> There is increasing evidence of a link between miRNA expression levels and the diagnosis of FTD<sup>8–12</sup> and ALS.<sup>11–22</sup> However, there are strong inconsistencies between the identified miRNA signatures in different studies. The examined cohorts are highly heterogeneous, most of them being sporadic or mixed cohorts of sporadic and genetic forms. Importantly, it is unclear which miRNAs are specific to a particular genetic mutation or might serve as biomarkers for several genetic forms. It is also uncertain whether miRNAs found in sporadic forms are differentially expressed in genetic forms. Furthermore, several of the published articles lacked an independent validation cohort, which also might have caused disparity between results. Finally, differences in lifestyle factors (e.g. diet, exercise, cognitive training) across the studied cohorts also influence the levels of blood-based biomarkers, and thus may contribute to non-reproducible results.<sup>23,24</sup> This absence of convergence among different studies so far hinders the use of miRNAs in clinical trials.

The present work aims at testing circulating miRNA signatures identified in the literature, using two independent homogeneous cohorts of patients and presymptomatic carriers: one focused on *C9orf72* expansion carriers and another comprising *GRN* mutation carriers. For that purpose, we selected all published studies that identified specific miRNAs isolated from plasma or serum as potential biomarkers of FTD and/or ALS. With a pre-registered study design, we investigated whether (1) miRNAs revealed in cohorts of sporadic patients (or in mixed cohorts with sporadic and genetic forms) may be biomarkers in *C9orf72* and/or *GRN* genetic forms, (2) miRNAs identified in a *C9orf72* cohort are validated in an independent *C9orf72* cohort, and (3) miRNAs discovered in a *C9orf72* cohort may be relevant in a *GRN* cohort.

We hypothesize that if a miRNA is a potential biomarker in a particular genetic form, it will be differentially

expressed (adjusted *p* value below 0.05) between controls and presymptomatic subjects, controls and patients, or presymptomatic subjects and patients in an independent cohort of subjects carrying that mutation. Moreover, we consider that a miRNA signature will constitute an acceptable biomarker if a logistic regression model (using these miRNAs as features) classifies subjects between clinical groups with an area under the ROC curve greater than 0.70.<sup>25</sup>

## Materials and Methods

This research was conducted according to the preregistration available in <https://osf.io/4pw8f>.

### Participants of the validation cohorts

Between 2011 and 2021, 135 individuals were recruited through the French research network on FTD/ALS (Inserm RBM02-59) and investigated with the same protocol, as previously described in detail.<sup>26</sup> All participants signed written informed consents. This study was approved by the Comité de Protection des Personnes CPP Ile-De-France VI (CPP 36–09/ID RCB 2008-A01376-49 and CPP 68–15/ID RCB 2015-A00856-43).

Two cohorts were studied. One cohort was focused on *C9orf72* mutation carriers, including 29 patients (20 FTD, 6 FTD/ALS and 3 ALS) and 17 carriers in the presymptomatic phase. Another cohort was focused on *GRN* mutation carriers, comprising 28 FTD patients and 30 presymptomatic carriers. The control group, shared between the two cohorts, was made up of 31 neurologically healthy individuals that did not carry any of these mutations. Table 1 displays the demographic characteristics of the studied cohorts.

Standardized interviews with family members, full neurological examinations, quantitative motor testing, and extensive neuropsychological tests measuring all cognitive domains were used to assess each participant's cognitive and clinical conditions. All subjects underwent blood tests, and collected samples were stored in the Paris Brain Institute (ICM) DNA and cell bank.

### Plasma preparation, miRNA sequencing and computation pipeline

Blood samples from all participants were collected on EDTA following standardized collection and handling procedures. The mean disease duration at sampling was 4.9 (SD 3.8) years in the *C9orf72* patients' group and 3.2 (SD 1.4) years in the *GRN* patients' group. All participants were in fasted state. At the ICM DNA and cell bank, all samples were centralized and processed in

**Table 1.** Demographic characteristics of the studied cohorts.

	<i>C9orf72</i>		<i>GRN</i> patients	<i>GRN</i> presymptomatic carriers	Controls
	<i>C9orf72</i> patients	presymptomatic carriers			
No.	29	17	28	30	31
Female, No. (%)	14 (48.3)	10 (58.8)	10 (35.7)	17 (56.7)	18 (58.1)
Age at inclusion, mean (SD)	66.2 (8.8)	51.7 (12.1)	62.9 (11.2)	42.5 (11)	47.1 (14.6)
Disease duration at inclusion, mean (SD)	4.9 (3.8)	–	3.2 (1.4)	–	–
<i>GRN</i> mutation (No.)	–	–	c.1157G > A (1), c.1231_1232dup (2), c.1252C > T (1), c.138 + 1G > A (1), c.1492G > T (1), c.1494_1498del (1), c.19 T > C (1), c.234_235del (1), c.25del (1), c.328C > T (2), c.380_381del (2), c.421_422del (2), c.559del (1), c.559delC (1), c.708 + 6_708 + 9del (1), c.745C > T (1), c.759_760del (1), c.768_769dup (1), c.87_90dup (1), c.900_901dup (1), c.907del (1), c.942C > A (1), Exon 1 del (2)	c.1201C > T (3), c.1231_1232dup (3), c.138 + 1G > A (1), c.1494_1498del (2), c.19 T > C (1), c.328C > T (1), c.361delG (1), c.380_381del (1), c.443_444del (1), c.675_676del (1), c.745C > T (3), c.768_769dup (2), c.813_816del (6), c.907del (2), c.988_989del (1), <i>GRN</i> del (1)	–

conformity with the same procedure. Blood samples were centrifuged at 2500 rpm for 10 min before plasma was extracted at room temperature. At a temperature of  $-80^{\circ}\text{C}$ , 1 ml aliquots were stored in polypropylene tubes.

Following the instructions provided by the manufacturer, miRNA extraction was carried out using a miR-Neasy Serum/Plasma Kit (Qiagen). We used 200  $\mu\text{l}$  of plasma that was progressively melted at  $4^{\circ}\text{C}$  and added directly to 1 ml of QIAzol solution. Using the QIAseq miRNA Library Kit (Qiagen) in accordance with the manufacturer's protocol, miRNAs were eluted in 14  $\mu\text{l}$  of water; 5  $\mu\text{l}$  were utilized to prepare the miRNA sequencing library. Targeting a minimum of 10 million mapped reads per sample, two independent batches of miRNA sequencing were performed on the Illumina NovaSeq 6000.

The quantification of miRNAs was carried out according to recommendations by Potla et al.<sup>27</sup> The quality of reads was assessed with FastQC (Andrews S. 2010, <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>).

Next, UMI-tools<sup>28</sup> and Cutadapt<sup>29</sup> were used to clean sequences and extract unique molecular identifiers (UMIs). Then, the resulting reads were aligned to the mature miRNA sequences from the miRBase (<https://www.mirbase.org>) database version 22.1, using Bowtie.<sup>30</sup> After that, the PCR duplicates were removed with UMI-tools. Finally, miRNA count tables were created with Samtools idxstats.<sup>31</sup>

## Selected studies

We aimed to find all papers that identified specific miRNAs extracted from human plasma or serum as potential biomarkers of FTD and/or ALS, excluding reviews and meta-analyses. We thus conducted the following search in PubMed (<https://pubmed.ncbi.nlm.nih.gov/>), on March 10, 2022:

(microRNA[Title] OR microRNAs[Title] OR miR[Title] OR miRNA[Title]) AND (serum[Title] OR circulating[Title] OR plasma[Title]) AND (ALS[Title] OR FTD[Title] OR amyotrophic[Title] OR frontotemporal[Title] OR (neurodegenerative[Title] AND (frontotemporal[Title/Abstract] OR amyotrophic[Title/Abstract]))) NOT mice[Title/Abstract] NOT mouse[Title/Abstract] NOT extracellular vesicles[Title] NOT review[PT] NOT meta-analysis[PT] NOT (comment[PT])

This search resulted in 19 journal articles. Two papers<sup>6,32</sup> were excluded because they were review studies, one<sup>33</sup> was discarded because it was focused on protein levels, and one<sup>34</sup> was excluded because it was focused on one microRNA from serum exosomes.

Our final selection therefore contained 15 articles. These selected papers, along with the studied diseases (FTD, ALS, or both), cohort types (sporadic, genetic, or mixed), cohort sizes, methods of analyses (qRT-PCR, microarrays, RNA-sequencing, or a combination), and the identified miRNA signatures are displayed in Table 2. We

**Table 2.** Selected studies investigating circulating microRNA expression (from serum or plasma) of patients with FTD or ALS. Columns indicate each reference, studied disease (FTD, ALS, or both), type of the analyzed cohort (sporadic, genetic, or mixed), number of patients in the discovery and replication (if available) cohorts, number of presymptomatic carriers included in the study, method of analysis (qRT-PCR, microarrays, RNA-sequencing, or a combination), and the identified miRNA signature.

Article	Disease	Cohort	Patients, No. (discovery/replication)	Presymptomatic carriers, No.	Method of analysis	MiRNA signature
Grasso et al., 2019 <sup>8</sup>	FTD	Sporadic	10/48 split of same cohort	–	qRT-PCR of 752 miRNAs	miR-663a, miR-502-3p, miR-206
Piscopo et al., 2018 <sup>9</sup>	FTD	Sporadic	54	–	qRT-PCR of 9 miRNAs linked with apoptosis	miR-127-3p
Denk et al., 2018 <sup>10</sup>	FTD	Sporadic	48	–	qRT-PCR of 96 miRNAs identified in preliminary study	let-7b-5p, let-7 g-5p, miR-106a-5p, miR-106b-5p, miR-18b-5p, miR-223-3p, miR-26a-5p, miR-26b-5p, miR-301a-3p, miR-30b-5p, miR-146a-5p, miR-15a-5p, miR-22-3p, miR-320a, miR-320b, miR-92a-3p, miR-1246
Kmetzsch et al., 2021 <sup>11</sup>	FTD, ALS	Genetic ( <i>C9orf72</i> )	22	45	RNA-sequencing of 2576 miRNAs	miR-34a-5p, miR-345-5p, miR-200c-3p, miR-10a-3p
Sheinerman et al., 2017 <sup>12</sup>	FTD, ALS	Unspecified	For each disease, 25/25 split of same cohort	–	qRT-PCR of 37 brain-enriched miRNAs	miR-9/let-7e, miR-7/miR-451, miR-335-5p/let-7e (FTD) and miR-206/miR-338-3p, miR-9/miR-129-3p, miR-335-5p/miR-338-3p (ALS)
Magen et al., 2021 <sup>13</sup>	ALS	Mixed sporadic and genetic ( <i>C9orf72</i> )	126/122 split of same cohort	–	RNA-sequencing of 125 miRNAs identified in longitudinal study	miR-181a-5p, miR-181b-5p
Soliman et al., 2021 <sup>14</sup>	ALS	Mixed sporadic and genetic (unspecified mutation)	30	–	qRT-PCR of 7 miRNAs involved in ALS	miR-206, miR-143-3p, miR-142-3p
Dobrowolny et al., 2021 <sup>15</sup>	ALS	Mixed sporadic and genetic (unspecified mutation)	13/23	–	RNA-sequencing followed by qRT-PCR	miR-151a-5p, miR-199a-5p, miR-423-3p
Raheja et al., 2018 <sup>16</sup>	ALS	Mixed sporadic and genetic ( <i>C9orf72</i> , <i>SOD1</i> )	23	–	qRT-PCR of 191 miRNAs identified on prior study	miR-29b-3p, miR-320c, miR-34a-5p, miR-29c-3p, miR-320a, miR-22-3p, miR-1, miR-133a-3p, miR-191-5p, miR-144-5p, miR-320b, miR-423-3p, miR-192-5p, miR-133b, miR-194-5p, miR-7-1-3p, miR-19a-3p, miR-425-5p, miR-145-5p, miR-144-3p
Waller et al., 2017 <sup>17</sup>	ALS	Sporadic	27/23	–	qRT-PCR of 750 miRNAs	miR-206, miR-143-3p, miR-374b-5p
Tasca et al., 2016 <sup>18</sup>	ALS	Sporadic	14	–	qRT-PCR of 9 muscle-specific,	miR-206, miR-133a, miR-133b, miR-27a

(Continued)

**Table 2** Continued.

Article	Disease	Cohort	Patients, No. (discovery/replication)	Presymptomatic carriers, No.	Method of analysis	MiRNA signature
Takahashi et al., 2015 <sup>19</sup>	ALS	Sporadic	16/48 split of same cohort	–	inflammatory, or angiogenic miRNAs Microarrays, followed by qRT-PCR of 9 miRNAs	miR-4649-5p, miR-4299
Freischmidt et al., 2015 <sup>20</sup>	ALS	Sporadic	18/20	–	Microarrays of 1733 miRNAs, followed by qRT-PCR of 2 miRNAs	miR-1234-3p, miR-1825
Freischmidt et al., 2014 <sup>21</sup>	ALS	Separate sporadic and genetic ( <i>SOD1</i> , <i>FUS</i> , <i>C9orf72</i> )	9/13 (genetic), 14 (sporadic)	18	Microarrays of 1733 miRNAs and qRT-PCR of 4 miRNAs	miR-4745-5p, miR-3665, miR-1915-3p, miR-4530
De Felice et al., 2014 <sup>22</sup>	ALS	Sporadic	10	–	qRT-PCR of 1 miRNA	miR-338-3p

note that three of these studies<sup>8–10</sup> identified three different miRNA signatures associated with FTD, 10 articles<sup>13–22</sup> pointed out 10 distinct miRNA signatures related to ALS, one investigation<sup>11</sup> revealed another miRNA signature for both FTD and ALS, and one study<sup>12</sup> found two separate signatures for FTD and ALS. In all, considering both FTD and ALS, we thus analyzed 16 miRNA signatures previously identified in the literature. Since *C9orf72* expansions can cause both diseases, all miRNA signatures were tested with our *C9orf72* cohort. However, since *GRN* mutations cause exclusively FTD, only five miRNA signatures (the ones associated with FTD) were tested with our *GRN* cohort.

Regarding the sizes of the identified miRNA signatures, most of the studies<sup>8,9,11,13–15,17–22</sup> pointed out four or less miRNAs, one article identified six miRNAs per disease,<sup>12</sup> and two investigations<sup>10,16</sup> proposed larger signatures containing respectively 17 or 20 miRNAs. When analyzing the intersections between the different signatures, we note that some miRNAs were identified by multiple studies, for instance miR-206,<sup>8,12,14,17,18</sup> but most miRNAs were found by a single study. When considering the union of the miRNA signatures from the selected articles, the set of miRNAs associated with either FTD or ALS is composed of 65 miRNAs, and the set of miRNAs associated only with FTD is composed of 30 miRNAs.

As previously mentioned in the description of our computation pipeline, we used the miRNA nomenclature from the miRBase database version 22.1. However, most of the selected articles were based on previous versions of the miRBase. We thus performed the following conversions for compatibility:

- miR-320a: miR-320a-5p plus miR-320a-3p
- miR-9: miR-9-5p plus miR-9-3p

- let-7e: let-7e-5p plus let-7e-3p
- miR-1: miR-1-5p plus miR-1-3p
- miR-133-a: miR-133-a-5p plus miR-133a-3p
- miR-27a: miR-27a-5p plus miR-27a-3p
- miR-7: miR-7-5p
- miR-451: miR-451a
- miR-129-3p: miR-129-1-3p

### Differential expression

Differential expression analyses were performed using the R package EdgeR.<sup>35</sup> After microRNA extraction, sequencing and quality control steps, our dataset contained the expression levels of 2656 miRNAs (denoted by  $i$ , corresponding to all miRNA sequences mapped in miRBase version 22.1) for each of the 135 subjects (represented by  $j$ ). First, we created two count matrices: one containing the miRNA counts from the *C9orf72* patients, presymptomatic subjects and controls, and another containing the miRNA counts from the *GRN* patients, presymptomatic individuals, and controls. Second, for each count matrix, we used generalized linear models (GLM) to fit a log-linear model to each miRNA, following a negative binomial distribution with mean  $\mu_{ij}$  and dispersion  $\phi_i$ :

$$\log_2 \mu_{ij} = x_j^T \beta_i$$

where  $x_j$  denotes the covariates describing sample  $j$  and  $\beta_i$  denotes the coefficients to be fitted for miRNA  $i$ . To control for possible age, sex, center and batch effects, we included them as covariates in the models, in addition to the clinical group (control, presymptomatic or patient). A trimmed mean of  $M$ -values<sup>36</sup> was used to normalize raw counts. Finally, after the log-linear models were fitted,

quasi-likelihood (QL) *F*-tests were used to identify the differentially expressed miRNAs.

Concretely, we tested each of the 65 miRNAs associated with either FTD or ALS in the literature, to identify which were differentially expressed between (a) controls versus *C9orf72* presymptomatic subjects, (b) controls versus *C9orf72* patients, and (c) *C9orf72* presymptomatic subjects versus *C9orf72* patients. Additionally, we tested the 30 miRNAs associated with only FTD in the literature, to highlight which were differentially expressed between (d) controls versus *GRN* presymptomatic subjects, (e) controls versus *GRN* patients, and (f) *GRN* presymptomatic subjects versus *GRN* patients.

All *p* values were 2-tailed, and the level of statistical significance was set at 0.05. The Benjamini-Hochberg<sup>37</sup> procedure was used to adjust *p* values for multiple testing. Additionally, we considered as suggestive, but not statistically significant, miRNAs with adjusted *p* values between 0.05 and 0.1. Our differential expression analyses resulted in two meta-signatures for each cohort: one meta-signature including miRNAs with adjusted *p* values below 0.05 (thus including only statistically significant miRNAs), and one containing miRNAs with adjusted *p* values below 0.1 (thus including both statistically significant and suggestive miRNAs).

## Binary classification

To test if the miRNA signatures described in the literature could discriminate between clinical groups, we trained L2-regularized logistic regression classifiers, using Python 3.8.5 with scikit-learn<sup>38</sup> 0.23.2. We first organized the miRNA expression data into six datasets, one for each relevant pairwise comparison: (a) controls versus *C9orf72* presymptomatic subjects, (b) controls versus *C9orf72* patients, (c) *C9orf72* presymptomatic subjects versus *C9orf72* patients, (d) controls versus *GRN* presymptomatic subjects, (e) controls versus *GRN* patients, and (f) *GRN* presymptomatic subjects versus *GRN* patients. A total of 18 classifiers were trained for each of the comparisons (a), (b) and (c): 16 classifiers used as features each of the miRNA signatures identified in the literature, and two were trained with meta-signatures containing the differentially expressed miRNAs identified in the univariate analyses (a), (b) and (c), respectively with adjusted *p* values lower than 0.05 and 0.1. In addition, seven classifiers were built for each of the comparisons (d), (e) and (f): five of them used as features each of the miRNA signatures associated with FTD in the literature, and two were trained with meta-signatures containing the differentially expressed miRNAs identified in the univariate analyses (d), (e) and (f), respectively with adjusted *p* values lower than 0.05 and 0.1.

Each logistic regression model was trained with a stratified nested 5-fold cross-validation strategy followed by bootstrapping to compute confidence intervals (CIs), as previously detailed.<sup>11</sup> The inner cross-validation loop was used for hyperparameter (L2 regularization coefficient) search. The outer cross-validation loop was used to assess each classifier's performance by computing the area under the ROC curve. Since a *k*-fold cross-validation does not provide an unbiased estimator of the variance,<sup>39</sup> it cannot be used to compute CIs. Therefore, we used 2000 bootstrap samples to compute empirical 90% CIs for the ROC AUC scores, by considering the 5th and 95th percentiles of the bootstrap distribution. A miRNA signature was considered an acceptable biomarker for a given comparison if the corresponding ROC AUC was above 0.70.

## Results

### Differential expression in the *C9orf72* cohort

The first analysis consisted of testing which of the 65 miRNAs identified in the literature as potential biomarkers of FTD and/or ALS were differentially expressed in our *C9orf72* cohort. For this analysis, we considered the miRNA counts obtained from sequencing plasma samples from the *C9orf72* patients, *C9orf72* presymptomatic subjects, and controls. After negative binomial generalized linear models were fitted for each miRNA, we thus performed 65 quasi-likelihood *F*-tests per pairwise comparison (controls versus *C9orf72* presymptomatic subjects, controls versus *C9orf72* patients, and *C9orf72* presymptomatic subjects versus *C9orf72* patients), adjusting for multiple comparisons using the Benjamini-Hochberg<sup>37</sup> method.

All differentially expressed miRNAs identified in this analysis are displayed in Table 3 (second to fourth columns). We can see that a considerable amount of miRNAs (35 of the 65 miRNAs identified in the literature) were significantly differentially expressed (adjusted *p* values smaller than 0.05) in at least one comparison: miR-34a-5p, miR-338-3p, miR-142-3p, miR-320a, miR-145-5p, miR-92a-3p, let-7 g-5p, miR-199a-5p, miR-206, miR-30b-5p, miR-191-5p, miR-27a, miR-320b, miR-143-3p, miR-1246, miR-223-3p, miR-144-3p, miR-451, miR-194-5p, miR-144-5p, miR-29b-3p, miR-29c-3p, miR-192-5p, miR-19a-3p, miR-502-3p, miR-15a-5p, miR-374b-5p, miR-7-1-3p, miR-320c, miR-106b-5p, miR-146a-5p, miR-133b, let-7b-5p, miR-345-5p, and miR-22-3p. Moreover, the following 9 miRNAs had a *p* value between 0.05 and 0.1, close to significance value: miR-151a-5p, miR-1234-3p, miR-26a-5p, miR-301a-3p, let-7e, miR-18b-5p, miR-106a-5p, miR-1915-3p, and miR-9.

**Table 3.** Differentially expressed miRNAs for at least one pairwise comparison between clinical groups, considering both cohorts. The \* indicates in which comparisons each miRNA was significantly differentially expressed (adjusted  $p$  values below 0.05), while the (+) denotes adjusted  $p$  values between 0.05 and 0.1.

miRNA	Controls vs. <i>C9orf72</i> presymptomatic subjects	Controls vs. <i>C9orf72</i> patients	<i>C9orf72</i> presymptomatic subjects vs. <i>C9orf72</i> patients	Controls vs. <i>GRN</i> presymptomatic subjects	Controls vs. <i>GRN</i> patients	<i>GRN</i> presymptomatic subjects vs. <i>GRN</i> patients
miR-34a-5p	*	*				
miR-338-3p	*		*			
miR-142-3p	*		*			
miR-320a	*		*			
miR-145-5p	*		*			
miR-92a-3p	*				(+)	
let-7g-5p	*					
miR-199a-5p	*		*			
miR-206	*	*	*			
miR-30b-5p	*		*			
miR-191-5p	*					
miR-27a	*		*			
miR-320b	*		*			
miR-143-3p	*		*			
miR-1246	*					
miR-223-3p	*		*			
miR-144-3p	*					
miR-451		*	(+)		*	(+)
miR-194-5p		*	*			
miR-144-5p		*	(+)			
miR-29b-3p		*	*			
miR-29c-3p		*	*			
miR-192-5p		*				
miR-19a-3p		*	*			
miR-502-3p		*			*	
miR-15a-5p		*			*	
miR-374b-5p	(+)		*			
miR-7-1-3p			*			
miR-320c	(+)		*			
miR-106b-5p			*		(+)	
miR-146a-5p	(+)		*			
miR-133b		(+)	*			
let-7b-5p	(+)		*		(+)	
miR-345-5p			*			
miR-22-3p			*			
miR-7					*	(+)
miR-18b-5p		(+)			*	
miR-151a-5p	(+)					
miR-1234-3p	(+)					
miR-26a-5p	(+)					
miR-301a-3p	(+)		(+)			
let-7e		(+)				
miR-106a-5p		(+)			(+)	
miR-1915-3p		(+)				
miR-9			(+)			

The complete output from the differential expression analyses in the *C9orf72* cohort, including log-fold changes indicating the intensity of underexpression or overexpression, as

well as computed  $p$  values, are displayed in Table S1. Expression heatmaps of the differentially expressed miRNAs are shown in Figure S1.



### Differential expression in the GRN cohort

The second analysis focused on identifying which of the 30 miRNAs linked with FTD in the literature were differentially expressed in our GRN cohort. For this experiment, we used the miRNA counts acquired by sequencing plasma samples from the GRN patients, GRN presymptomatic individuals, and controls. Once negative binomial generalized linear models were fitted for each miRNA, we conducted 30 quasi-likelihood F-tests per pairwise comparison (controls versus GRN presymptomatic subjects, controls versus GRN patients, and GRN presymptomatic subjects versus GRN patients), adjusting for multiple comparisons using the Benjamini-Hochberg<sup>37</sup> method.

Table 3 (fifth to seventh columns) shows all differentially expressed miRNAs identified in this experiment. Contrary to what was observed with the *C9orf72* cohort, we note that only a small proportion of miRNAs (5 of the 30 miRNAs identified in the literature) were significantly differentially expressed (adjusted *p* values lower than 0.05), all of them when comparing controls and GRN patients: miR-451, miR-15a-5p, miR-502-3p, miR-7, and miR-18b-5p. Additionally, 4 miRNAs had a *p* value close to significance value, between 0.05 and 0.1: miR-106a-5p, miR-92a-3p, miR-106b-5p, and let-7b-5p.

Table S2 summarizes the complete results of the differential expression experiments in the GRN cohort, including log-fold changes reflecting the degree of underexpression or overexpression of each miRNA in each pairwise comparison, and the calculated *p* values. Figure S1 displays expression heatmaps of the differentially expressed miRNAs.

Finally, Table 3 also allows comparing the results obtained with the *C9orf72* and the GRN cohorts. Remarkably, the three comparisons involving the *C9orf72* cohort revealed significantly differentially expressed miRNAs, but that was the case for only one comparison involving the GRN cohort (controls versus GRN patients). We also note that only a small minority of miRNAs (3) was significantly differentially expressed in both cohorts.

Taken together, these results offer evidence for the potential contribution of miRNAs identified in previous studies as biomarkers of *C9orf72*-associated disease, but not GRN-associated disease.

### Binary classification in the C9orf72 cohort

The first set of binary classification experiments focused on the *C9orf72* cohort. We trained 18 logistic regression classifiers for each pairwise comparison between clinical groups (controls versus *C9orf72* presymptomatic subjects, controls versus *C9orf72* patients, and *C9orf72*

presymptomatic subjects versus *C9orf72* patients). For each of the 18 classifiers, we used as features the expression levels of distinct sets of miRNAs: the 16 miRNA signatures identified in the previously published studies about FTD and ALS, and two meta-signatures obtained from our differential expression analyses with the *C9orf72* cohort (one containing miRNAs with adjusted *p* value <0.05, and another including miRNAs with adjusted *p* value <0.1).

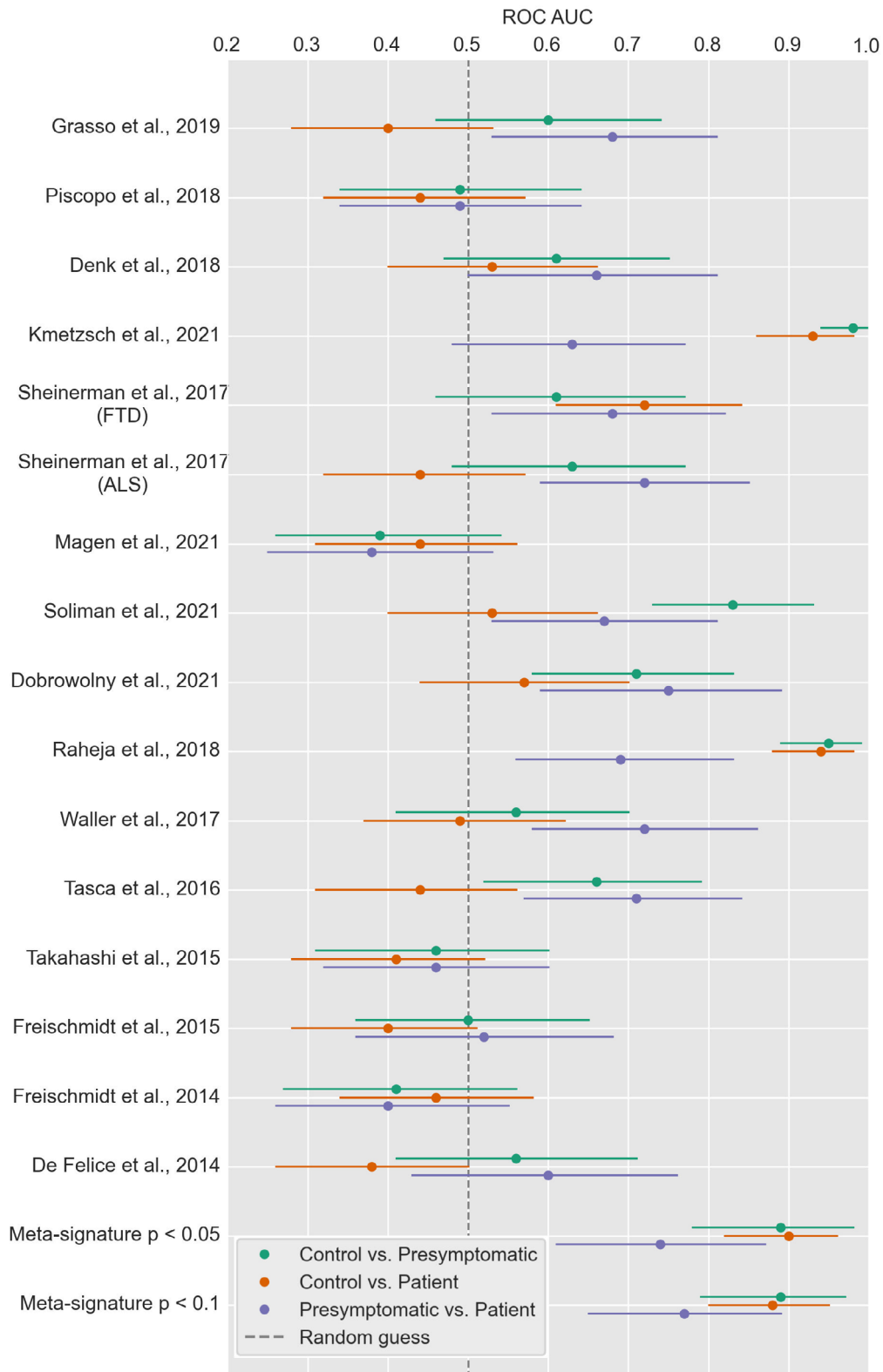
Figure 1 displays the areas under the ROC curves obtained by each of the 18 logistic regression classifiers in the three pairwise comparisons, as well as the empirical 90% confidence intervals. We observe that more than half of the classifiers (10 of the 18) achieved a ROC AUC greater than 0.70 in at least one comparison: those using the miRNA signatures from Kmetzsch et al., 2021,<sup>11</sup> Sheinerman et al., 2017<sup>12</sup> (FTD), Sheinerman et al., 2017<sup>12</sup> (ALS), Soliman et al., 2021,<sup>14</sup> Dobrowolny et al., 2021,<sup>15</sup> Raheja et al., 2018,<sup>16</sup> Waller et al., 2017,<sup>17</sup> Tasca et al., 2016,<sup>18</sup> and the two meta-signatures from our differential expression analyses. The miRNA signatures with the largest ROC AUC were from Kmetzsch et al., 2021<sup>11</sup> (0.98 for controls versus presymptomatic subjects), Raheja et al., 2018<sup>16</sup> (0.94 for controls versus patients), and the meta-signature with *p* < 0.1 (0.77 for presymptomatic subjects versus patients).

As well as in the differential expression analyses, these findings provide support for the potential use of several miRNAs identified in previous studies as biomarkers of *C9orf72*-associated FTD and ALS.

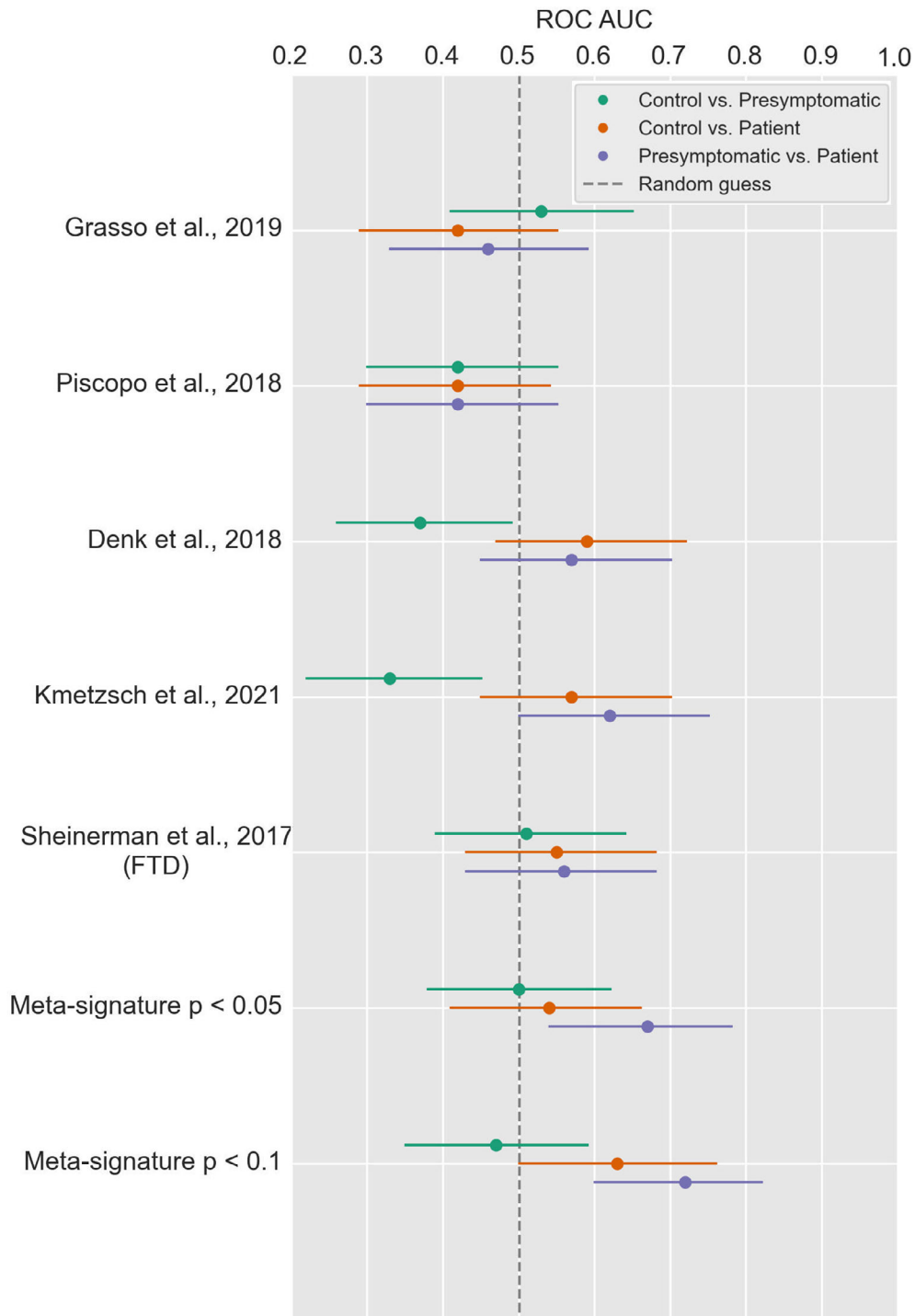
### Binary classification in the GRN cohort

The second set of binary classification experiments consisted of training seven logistic regression classifiers for each pairwise comparison in the GRN cohort (controls versus GRN presymptomatic subjects, controls versus GRN patients, and GRN presymptomatic subjects versus GRN patients). Each of the seven classifiers was trained with a different set of features: the expression levels of the five miRNA signatures linked with FTD in previously published studies, and two meta-signatures obtained in our differential expression analyses with the GRN cohort (one consisting of miRNAs with an adjusted *p* value smaller than 0.05, and another comprising miRNAs with an adjusted *p* value smaller than 0.1).

The areas under the ROC curves and empirical 90% confidence intervals for each of the seven logistic regression classifiers, in the three pairwise comparison involving the GRN cohort, are shown in Figure 2. Strikingly, only one ROC AUC was (slightly) greater than 0.70: when classifying GRN presymptomatic subjects and GRN patients, using the meta-signature from our differential



**Figure 1.** Area under the ROC curve results when classifying groups from the *C9orf72* cohort. The solid circles indicate the areas under the ROC curves obtained for each pairwise comparison using 18 different miRNA signatures. The whiskers denote empirical 90% confidence intervals obtained with 2000 bootstrap samples.



**Figure 2.** Area under the ROC curve results when classifying groups from the *GRN* cohort. The solid circles indicate the areas under the ROC curves obtained for each pairwise comparison using seven different miRNA signatures. The whiskers denote empirical 90% confidence intervals obtained with 2000 bootstrap samples.

expression analysis comprising miRNAs with *p* value lower than 0.1. The miRNA signatures with the largest ROC AUC were from Grasso et al., 2019<sup>8</sup> (0.53 for

controls versus presymptomatic subjects), and the meta-signature with *p* < 0.1 (0.63 for controls versus patients, and 0.72 for presymptomatic subjects versus patients).

It is noteworthy that these results are consistent with our differential expression analyses with the *GRN* cohort, offering evidence that most miRNAs identified in previous studies about FTD are not useful biomarkers of *GRN*-associated disease.

## Discussion

The goal of this study was to assess all circulating miRNA signatures previously published in the literature as possible biomarkers of FTD and/or ALS, by testing them in two separate homogenous cohorts of *C9orf72* and *GRN* mutation carriers, comprising patients and presymptomatic subjects. The results of this work demonstrate that (1) several miRNAs identified in sporadic or mixed FTD/ALS cohorts could potentially be used as biomarkers of *C9orf72* disease; (2) some miRNAs revealed in a *C9orf72* cohort are validated in an independent *C9orf72* cohort; and (3) most miRNAs associated with FTD in sporadic or mixed cohorts, or in a cohort of *C9orf72* mutation carriers, are not relevant biomarkers of *GRN* disease.

First, differential expression results (Table 3) showed that more than half (35/65) of the miRNAs linked with FTD and/or ALS in the literature were significantly differentially expressed in the *C9orf72* cohort. Remarkably, only four of the 15 selected studies included *C9orf72* mutation carriers,<sup>11,13,16,21</sup> three of which focused exclusively on ALS.<sup>13,16,21</sup> Therefore, these outcomes reveal similarities in miRNA dysregulation between individuals with sporadic forms of FTD/ALS and *C9orf72*-associated disease. Classification results with the *C9orf72* cohort (Fig. 1) also corroborate these findings, since half of the examined miRNA signatures (8/16) yielded at least one pairwise comparison with acceptable performance, and all comparisons employing the meta-signatures had acceptable performance. We considered as an acceptable biomarker any signature with a ROC AUC above 0.70, in accordance with the recommendation by Mandrekar.<sup>25</sup> This means that such a signature has some discriminatory power and is worthy of further exploration. It does not mean that the signature is suitable for clinical use.

Next, we observed that a miRNA signature previously identified in a homogeneous *C9orf72* cohort<sup>11</sup> and another one revealed in a mixed cohort of sporadic and familial ALS<sup>16</sup> displayed an outstanding result (ROC AUC above 0.90) when classifying controls versus *C9orf72* presymptomatic subjects and controls versus *C9orf72* patients (Fig. 1). These two signatures have in common the presence of miR-34a-5p, which has the smallest adjusted *p* value in the differential expression analyses regarding these comparisons (Table S1, respectively

$p = 2.42E-08$  and  $p = 5.06E-06$ ). In contrast, the performance of both of these signatures classifying *C9orf72* presymptomatic individuals from patients was unsatisfactory. Indeed, neither of them contained miR-206, which is the most differentially expressed miRNA in this comparison (Table S1,  $p = 9.04E-05$ ). The overexpression of miR-206 in ALS patients had already been evidenced,<sup>40</sup> and the results of the present work extend this association also to *C9orf72*-disease. Nevertheless, even using the expression levels of miR-206, the classification of *C9orf72* presymptomatic subjects versus patients led to lower performances than comparisons involving the control group: the highest ROC AUC was 0.77, using the meta-signature with  $p < 0.1$ .

Finally, our results with the *GRN* cohort suggest that previously identified miRNAs have a weaker correlation with disease diagnosis and progression in this genetic form. Only a small proportion (5/30) of the miRNAs associated with FTD in previous studies was significantly differentially expressed in the *GRN* cohort (Table 3), and not a single miRNA was differentially expressed between controls and presymptomatic *GRN* carriers. Regarding the classification experiments, none of the studied miRNA signatures in the *GRN* cohort exhibited an acceptable performance (Fig. 2), and the only ROC AUC slightly above 0.70 was obtained when classifying *GRN* presymptomatic carriers and patients using the largest meta-signature (miRNAs with *p* value  $< 0.1$ ). One should note that none of the previous studies included *GRN* participants. Thus, our results demonstrate that miRNAs associated with sporadic FTD or genetic FTD due to *C9orf72* are not relevant for *GRN*-associated disease. Further studies are needed to determine if other miRNAs, not analyzed in the present paper, are useful in *GRN*-associated disease.

Validation studies using independent datasets, such as this one, are crucial to assess the utility of biomarker candidates, fostering research rigor and reproducibility. Notably, we carefully defined our research questions and analysis plan before data analysis, and preregistered our study. Preregistration has the strong benefit of leaving no flexibility for changes in analytical decisions after observing the data, which has been highlighted as a major source of false discoveries and replication failure.<sup>41</sup>

The main limitation of this work is the size of the studied cohorts, particularly the small group of *C9orf72* presymptomatic carriers (17) in comparison with the other groups, due to the rarity of genetic FTD. Additionally, due to the low number of *C9orf72* patients with different phenotypes (20 FTD, 6 FTD/ALS and 3 ALS), no conclusions can be drawn concerning the relationship between miRNAs and different disease manifestations. Moreover, this study did not investigate the influence of

lifestyle factors in microRNA levels. Future work will explore the combination of circulating microRNAs with other biomarkers, such as gray matter volume,<sup>5</sup> white matter integrity,<sup>42</sup> and neurofilament light chain level.<sup>43</sup> Multimodality will be crucial to accurately assess progression in *GRN*-associated FTD, and will likely improve the understanding of *C9orf72*-associated disease.

In summary, the present work revealed that most miRNAs previously identified in sporadic or mixed FTD/ALS cohorts are potential biomarkers of *C9orf72*-associated FTD/ALS, but not of *GRN*-associated FTD. Longitudinal studies are needed to confirm our findings and to determine miRNAs expression levels changes throughout disease progression, before circulating miRNAs can be used to assess treatment outcomes in *C9orf72*-associated disease clinical trials.

## Acknowledgements

We thank Justine Guégan for microRNA raw reads to counts pipeline handling. The study was conducted with the support of the Centre d'Investigation Clinique Neurosciences (CIC 1422), GH Pitié Salpêtrière, Paris, and the Centre pour l'Acquisition et le Traitement des Images platform. The research leading to these results has received funding from the program "Investissements d'avenir" ANR-10-IAIHU-06, from the French government under management of Agence Nationale de la Recherche as part of the "Investissements d'avenir" program, reference ANR-19-P3IA-0001 (PRAIRIE 3IA Institute), from Agence Nationale de la Recherche/DGOS (project ANR-PRTS PREV-DEMALS, grant number ANR-14-CE15-0016-07, promotion Assistance Publique – Hôpitaux de Paris), from the Programme Hospitalier de Recherche Clinique (PHRC) Predict-PGRN (to ILB, promotion by Assistance Publique – Hôpitaux de Paris), from Fondation Plan Alzheimer (to ILB), from the Inria Project Lab Program (project Neuromarkers), from Fondation Vaincre Alzheimer FR-17035 and from the Institut Français de Bioinformatique (ANR-11-INSB-0013).

## Conflict of Interest

The authors declare no conflict of interest.

## Author Contributions

Virgilio Kmetzsch had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Conception and design of the study: all authors. Data acquisition, analysis and interpretation: all authors. Statistical analysis: VK. Drafting of the manuscript: VK. Revising the manuscript: all authors. The

French research network on FTD/ALS contributed to data acquisition: Sophie Auriacombe (CHU Pellegrin, Bordeaux), Serge Belliard (CHU Rennes), Frédéric Blanc (Hôpitaux Civils, Strasbourg), Claire Boutoleau-Bretonnière (CHU Laennec, Nantes), Alexis Brice (Hôpital Pitié-Salpêtrière, Paris), Agnès Camuzat (ICM, Paris), Mathieu Ceccaldi (CHU La Timone, Marseille), Philippe Couratier (CHU Limoges), Vincent Deramecourt (CHU Roger Salengro, Lille), Mira Didic (CHU La Timone, Marseille), Charles Duyckaerts (Hôpital Pitié-Salpêtrière, Paris), Frédérique Etcharry-Bouyx (CHU Angers), Maité Formaglio (CHU Lyon), Véronique Golfier (CHU Rennes), Didier Hannequin (CHU Charles Nicolle, Rouen), Lucette Lacomblez (Hôpital Pitié-Salpêtrière, Paris), Isabelle Le Ber (Hôpital Pitié-Salpêtrière, Paris), Bernard-François Michel (CH Sainte-Marguerite, Marseille), Jérémie Pariente (CHU Ranguel, Toulouse), Florence Pasquier (CHU Lille), Daisy Rinaldi (CHU Pitié-Salpêtrière, Paris), Adeline Rollin-Sillaire (CHU Roger Salengro, Lille), Mathilde Sauvé (CHU Grenoble Alpes), François Sellal (CH Colmar), Christel Thauvin-Robinet (CHU Dijon), Catherine Thomas-Anterion (CH Plein-Ciel, Lyon), Martine Vercelletto (CHU Laennec, Nantes) and David Wallon (CHU Charles Nicolle Rouen).

## REFERENCES

1. Neary D, Snowden J, Mann D. Frontotemporal dementia. *Lancet Neurol.* 2005;4(11):771-780. doi:10.1016/S1474-4422(05)70223-4
2. DeJesus HM, Mackenzie IR, Boeve BF, et al. Expanded GGGGCC hexanucleotide repeat in noncoding region of C9ORF72 causes chromosome 9p-linked FTD and ALS. *Neuron.* 2011;72(2):245-256. doi:10.1016/j.neuron.2011.09.011
3. Renton AE, Majounie E, Waite A, et al. A hexanucleotide repeat expansion in C9ORF72 is the cause of chromosome 9p21-linked ALS-FTD. *Neuron.* 2011;72(2):257-268. doi:10.1016/j.neuron.2011.09.010
4. Pasinelli P, Brown RH. Molecular biology of amyotrophic lateral sclerosis: insights from genetics. *Nat Rev Neurosci.* 2006;7(9):710-723. doi:10.1038/nrn1971
5. Rohrer JD, Nicholas JM, Cash DM, et al. Presymptomatic cognitive and neuroanatomical changes in genetic frontotemporal dementia in the genetic frontotemporal dementia initiative (GENFI) study: a cross-sectional analysis. *Lancet Neurol.* 2015;14(3):253-262. doi:10.1016/S1474-4422(14)70324-2
6. Grasso M, Piscopo P, Crestini A, Confalonni A, Denti MA. Circulating microRNAs in neurodegenerative diseases. In: Igaz P, ed. *Circulating MicroRNAs in Disease Diagnostics and their Potential Biological Relevance.* *Experientia Supplementum.* Springer; 2015:151-169. doi:10.1007/978-3-0348-0955-9\_7

7. Huntzinger E, Izaurralde E. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. *Nat Rev Genet.* 2011;12(2):99-110. doi:10.1038/nrg2936
8. Grasso M, Piscopo P, Talarico G, et al. Plasma microRNA profiling distinguishes patients with frontotemporal dementia from healthy subjects. *Neurobiol Aging.* 2019;84:240.e1-240.e12. doi:10.1016/j.neurobiolaging.2019.01.024
9. Piscopo P, Grasso M, Puopolo M, et al. Circulating miR-127-3p as a potential biomarker for differential diagnosis in frontotemporal dementia. *J Alzheimers Dis.* 2018;65(2):455-464. doi:10.3233/JAD-180364
10. Denk J, Oberhauser F, Kornhuber J, et al. Specific serum and CSF microRNA profiles distinguish sporadic behavioural variant of frontotemporal dementia compared with Alzheimer patients and cognitively healthy controls. *PloS One.* 2018;13(5):e0197329. doi:10.1371/journal.pone.0197329
11. Kmetzsch V, Anquetil V, Saracino D, et al. Plasma microRNA signature in presymptomatic and symptomatic subjects with C9orf72-associated frontotemporal dementia and amyotrophic lateral sclerosis. *J Neurol Neurosurg Psychiatry.* 2021;92(5):485-493. doi:10.1136/jnnp-2020-324647
12. Sheinerman KS, Toledo JB, Tsvinsky VG, et al. Circulating brain-enriched microRNAs as novel biomarkers for detection and differentiation of neurodegenerative diseases. *Alzheimers Res Ther.* 2017;9(1):89. doi:10.1186/s13195-017-0316-0
13. Magen I, Yacovzada NS, Yanowski E, et al. Circulating miR-181 is a prognostic biomarker for amyotrophic lateral sclerosis. *Nat Neurosci.* 2021;24(11):1534-1541. doi:10.1038/s41593-021-00936-z
14. Soliman R, Mousa NO, Rashed HR, et al. Assessment of diagnostic potential of some circulating microRNAs in amyotrophic lateral sclerosis patients, an Egyptian study. *Clin Neurol Neurosurg.* 2021;208:106883. doi:10.1016/j.clineuro.2021.106883
15. Dobrowolny G, Martone J, Lepore E, et al. A longitudinal study defined circulating microRNAs as reliable biomarkers for disease prognosis and progression in ALS human patients. *Cell Death Discov.* 2021;7:4. doi:10.1038/s41420-020-00397-6
16. Raheja R, Regev K, Healy BC, et al. Correlating serum microRNAs and clinical parameters in amyotrophic lateral sclerosis. *Muscle Nerve.* 2018;58(2):261-269. doi:10.1002/mus.26106
17. Waller R, Goodall EF, Milo M, et al. Serum miRNAs miR-206, 143-3p and 374b-5p as potential biomarkers for amyotrophic lateral sclerosis (ALS). *Neurobiol Aging.* 2017;55:123-131. doi:10.1016/j.neurobiolaging.2017.03.027
18. Tasca E, Pegoraro V, Merico A, Angelini C. Circulating microRNAs as biomarkers of muscle differentiation and atrophy in ALS. *Clin Neuropathol.* 2016;35(1):22-30. doi:10.5414/NP300889
19. Takahashi I, Hama Y, Matsushima M, et al. Identification of plasma microRNAs as a biomarker of sporadic amyotrophic lateral sclerosis. *Mol Brain.* 2015;8(1):67. doi:10.1186/s13041-015-0161-7
20. Freischmidt A, Müller K, Zondler L, et al. Serum microRNAs in sporadic amyotrophic lateral sclerosis. *Neurobiol Aging.* 2015;36(9):2660.e15-2660.e20. doi:10.1016/j.neurobiolaging.2015.06.003
21. Freischmidt A, Müller K, Zondler L, et al. Serum microRNAs in patients with genetic amyotrophic lateral sclerosis and pre-manifest mutation carriers. *Brain J Neurol.* 2014;137(Pt 11):2938-2950. doi:10.1093/brain/awu249
22. De Felice B, Annunziata A, Fiorentino G, et al. miR-338-3p is over-expressed in blood, CFS, serum and spinal cord from sporadic amyotrophic lateral sclerosis patients. *Neurogenetics.* 2014;15(4):243-253. doi:10.1007/s10048-014-0420-2
23. Simrén J, Gustafson DR. A blood-based biomarker of cognitive decline and interaction with lifestyle. *JAMA Netw Open.* 2022;5(3):e223602. doi:10.1001/jamanetworkopen.2022.3602
24. Henriksen K, O'Bryant SE, Hampel H, et al. The future of blood-based biomarkers for Alzheimer's disease. *Alzheimers Dement J Alzheimers Assoc.* 2014;10(1):115-131. doi:10.1016/j.jalz.2013.01.013
25. Mandrekar JN. Receiver operating characteristic curve in diagnostic test assessment. *J Thorac Oncol.* 2010;5(9):1315-1316. doi:10.1097/JTO.0b013e3181ec173d
26. Le Ber I, Guedj E, Gabelle A, et al. Demographic, neurological and behavioural characteristics and brain perfusion SPECT in frontal variant of frontotemporal dementia. *Brain J Neurol.* 2006;129(Pt 11):3051-3065. doi:10.1093/brain/awl288
27. Potla P, Ali SA, Kapoor M. A bioinformatics approach to microRNA-sequencing analysis. *Osteoarthritis Cartilage.* 2021;3(1):100131. doi:10.1016/j.ocarto.2020.100131
28. Smith T, Heger A, Sudbery I. UMI-tools: modeling sequencing errors in unique molecular identifiers to improve quantification accuracy. *Genome Res.* 2017;27(3):491-499. doi:10.1101/gr.209601.116
29. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnetjournal.* 2011;17(1):10-12. doi:10.14806/ej.17.1.200
30. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25. doi:10.1186/gb-2009-10-3-r25
31. Li H, Handsaker B, Wysoker A, et al. The sequence alignment/map format and SAMtools. *Bioinforma Oxf Engl.* 2009;25(16):2078-2079. doi:10.1093/bioinformatics/btp352
32. Brennan S, Keon M, Liu B, Su Z, Saksena NK. Panoramic visualization of circulating MicroRNAs across

neurodegenerative diseases in humans. *Mol Neurobiol*. 2019;56(11):7380-7407. doi:10.1007/s12035-019-1615-1

33. Freischmidt A, Goswami A, Limm K, et al. A serum microRNA sequence reveals fragile X protein pathology in amyotrophic lateral sclerosis. *Brain*. 2021;144(4):1214-1229. doi:10.1093/brain/awab018

34. Xu Q, Zhao Y, Zhou X, Luan J, Cui Y, Han J. Comparison of the extraction and determination of serum exosome and miRNA in serum and the detection of miR-27a-3p in serum exosome of ALS patients. *Intractable Rare Dis Res*. 2018;7(1):13-18. doi:10.5582/irdr.2017.01091

35. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinforma Oxf Engl*. 2010;26(1):139-140. doi:10.1093/bioinformatics/btp616

36. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol*. 2010;11(3):R25. doi:10.1186/gb-2010-11-3-r25

37. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Methodol*. 1995;57(1):289-300.

38. Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: machine learning in python. *J Mach Learn Res*. 2011;12 (Oct):2825-2830.

39. Bengio Y, Grandvalet Y. No unbiased estimator of the variance of K-fold cross-validation. *J Mach Learn Res*. 2004;5(Sep):1089-1105.

40. Toivonen JM, Manzano R, Oliván S, Zaragoza P, García-Redondo A, Osta R. MicroRNA-206: a potential circulating biomarker candidate for amyotrophic lateral sclerosis. *PLoS One*. 2014;9(2):e89065. doi:10.1371/journal.pone.0089065

41. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution. *Proc Natl Acad Sci USA*. 2018;115(11):2600-2606. doi:10.1073/pnas.1708274114

42. Bertrand A, Wen J, Rinaldi D, et al. Early cognitive, structural, and microstructural changes in Presymptomatic C9orf72 carriers younger than 40 years. *JAMA Neurol*. 2018;75(2):236-245. doi:10.1001/jamaneurol.2017.4266

43. Saracino D, Dorgham K, Camuzat A, et al. Plasma NFL levels and longitudinal change rates in C9orf72 and GRN-associated diseases: from tailored references to clinical

applications. *J Neurol Neurosurg Psychiatry*. 2021;92 (12):1278-1288. doi:10.1136/jnnp-2021-326914

## Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Figure S1** Expression heatmaps of differentially expressed microRNAs. Rows represent microRNAs and columns represent individuals ordered by clinical status (control, presymptomatic, and patient). The log2 expression values of microRNAs are standardized (mean of 0 and standard deviation of 1), and z-scores are indicated by colors (blue indicates underexpression and red indicates overexpression). (A) All 35 differentially expressed microRNAs identified in the *C9orf72* cohort. (B) Zoom over five of the most differentially expressed microRNAs identified in the *C9orf72* cohort. (C) All five differentially expressed microRNAs identified in the *GRN* cohort.

**Table S1** Complete output from differential expression analyses in the *C9orf72* cohort, for each pairwise comparison between the clinical groups. The columns show the 30 studied miRNAs, the log-fold change when comparing the clinical groups, the unadjusted *p* values, and finally the adjusted *p* values after Benjamini-Hochberg. For each pairwise comparison, a positive log-fold change means that the miRNA is overexpressed in the first group. Controls (*n* = 31), *C9orf72* presymptomatic subjects (*n* = 17), and *C9orf72* patients (*n* = 29). Adjusted *p* values lower than 0.05 are shown in bold.

**Table S2** Complete output from differential expression analyses in the *GRN* cohort, for each pairwise comparison between clinical groups. The columns show the 30 studied miRNAs, the log-fold change when comparing the clinical groups, the unadjusted *p* values, and finally the adjusted *p* values after Benjamini-Hochberg. For each pairwise comparison, a positive log-fold change means that the miRNA is overexpressed in the first group. Controls (*n* = 31), *GRN* presymptomatic subjects (*n* = 30), *GRN* patients (*n* = 28). Adjusted *p* values lower than 0.05 are shown in bold.