Supplement 1

Title: miRNA Biomarkers Diagnose Amyotrophic Lateral Sclerosis in Circulating Blood

Authors: Rachael Anne Dunlop[1] • Paul Alan Cox[1] • Paul Mehta[2] • Elijah W. Stommel[3] • Sandra Anne Banack[1*].

Affiliations: [1]Brain Chemistry Labs, Jackson, WY, 83001, USA. [2]Office of Innovation and Analytics, Agency for Toxic Substances and Disease Registry, Centers for Disease Control and Prevention, Atlanta, GA, 30033, USA. [3]Department of Neurology, Dartmouth-Hitchcock Medical Center, Lebanon, NH, 03756, USA.

*Corresponding author: sandra@brainlabs.org

eMethods: Development of this 8-miRNA ALS diagnostic test involved: (1) extraction, immunoassay purification, and characterization of neural-enriched extracellular vesicles (NEE); (2) a feasibility study in four healthy controls to determine, using next generation sequencing (NGS), the quantity and quality of miRNA found in NEE and whether the miRNA within two fractions obtained during the NEE extraction process were differentially expressed; (3) identification, using NGS, of all small RNA found in the blood plasma of 10 ALS patients and 10 healthy controls; (4) evaluation of 101 differentially expressed miRNA to select those with the greatest separation between groups for further testing; (5) relative-quantification using qPCR of 34 candidate miRNA potentially useful for ALS diagnoses using 20 total ALS patients and 20 controls divided into two separate cohorts; (6) identification of an 8-miRNA ALS fingerprint (miR-151a-3p, miR-151a-5p, miR-4454, miR-146a-5p, miR-10b-5p, miR-199a-3p, miR-199a-5p, miR-29b-3p) to diagnose ALS; (7) qPCR method optimization; (8) cohort verification of fold-regulation patterns from the 8-miRNA, NEE-extracted, ALS fingerprint totalling 449 individuals; (9) development of machine learning algorithms to determine classification efficiencies of the 8-miRNA ALS fingerprint; (10) disproving the hypothesis that the 8-miRNA ALS fingerprint was representative of general neurodegeneration and not specific to ALS by comparing the same fingerprint against other neurological diseases (Primary Lateral Sclerosis and Parkinson's Disease); (11) testing and disproving the hypothesis that the miRNA, identified and quantified following NEE extraction, were unique to NEE and, therefore, could not be extracted directly from circulating blood plasma to achieve similar ALS diagnostic classification accuracy; (12) cohort verification groups resulting in 788 individuals tested in a total of five cohorts [1-5, current manuscript].

Criteria for patient eligibility, outcomes, and predictors were identical for the original method which included an NEE extraction step before extracting miRNA and for the current method which used miRNA extracted from circulating plasma without the NEE extraction step. The overall sample size in this paper was determined by achieving statistical accuracy for AUC calculations which require a sample size greater than 300 and by the availability of ALS blood plasma (a rare disease) from previously untested patients. The number of healthy individuals was chosen to match the number of ALS patients within each cohort with an emphasis on obtaining comparable age and sex individuals. All individuals were included in the final logistic regression model and cross-validation was used to avoid data partitioning. Internal validation was tested by creating a data-subset, run against the same model created with the entire dataset, to check for comparable confusion matrix parameters and to check if the same miRNA were selected for model inclusion. A minimum lambda value was chosen to maximize the model accuracy, however, a lambda of 1se produced identical results due to the robust nature of the model. The model threshold was set and reported at 50% prediction probability, however, increasing the threshold cutoff to 52% probability improved specificity and positive predicted values to 94% without changing sensitivity or negative predictive values.

RNA Extraction.

Aliquots of control and ALS plasma were removed from a -80˚C freezer, thawed at room temperature, then 200 µL were aliquoted into 1.5 mL LoBind® Tubes (catalogue number 022431021**,** Eppendorf). Total RNA was extracted using the Qiagen miRNeasy Serum/Plasma Kit (50) (catalogue number 217184) according to the manufacturer's instructions. Briefly, 1 mL QIAzol Lysis reagent (Qiagen, catalogue number 79306) was prepared with 1 µL of the synthetic spike-ins UniSp2, 4 and 5 (present at 2.0, 0.02, and 0.00002 fmol/µL, respectively) from the Qiagen RNA Spike-In Kit, For RT (catalogue number 339390). 1 mL of this preparation was added to each sample and homogenized by pipetting up and down five times. Samples were incubated at room temperature for 5 mins, then 200 µL chloroform were added to each sample, and the tubes shaken vigorously for 15 secs. Samples were incubated at room temperature for 3 mins, then centrifuged at 12,000 x g for 15 min at 4˚C, to generate three distinct layers. The top aqueous layer containing RNA was carefully removed (without touching the middle layer containing DNA) and placed in a new clean 2 mL LoBind® tube (catalogue number 022431048, Eppendorf). To this, we added 1.5 volumes of 100% ethanol (to 740 µL RNA we added 1110 µL 100% ethanol). We mixed samples thoroughly by pipetting up and down, five times. To each RNeasy MinElute Spin Column, we added 700 µL of sample, then centrifuged the tube at 10,000 x g for 15 secs at room temperature, then discarded the eluent. This step was repeated until all the sample had been added to the spin

column (three spins). We pipetted 700 µL Buffer RWT into each column, then spun at 10,000 x g for 15 secs at room temperature. The eluent was discarded. Next, we added 500 µL Buffer RPE to the spin-column then centrifuged at 10,000 x g for 15 secs at room temperature and discarded the eluent. We then added 500 µL of 80% ethanol to the spin column and spun for 2 min at 10,000 x g at room temperature, and the eluent was discarded. The spin-column was placed in a new 2 mL collection tube and centrifuged at 17,000 x g for 5 min with the spin column lid open at room temperature to dry the spin column. Finally, the spin column was placed in a clean 1.5 mL LoBind® tube, and 15 µL nuclease-free water was added directly to the membrane. The spin columns were spun with the lid closed for 1 min at 17,000 x g at room temperature to elute the RNA. The RNA was frozen at -80˚C until required for cDNA synthesis. Due to the highly valuable nature of the samples, only a single RNA extraction was performed per sample.

eTable1: Catalogue and GeneGlobe IDs for Qiagen miRCURY LNA PCR Assays.

		Qiagen miRCURY LNA miRNA PCR Assays catalogue number: 339306	
		Name	GeneGlobe ID
	Targets & reference genes	hsa-miR-10b-5p	YP00205637
		hsa-miR-4454	YP02114119
		hsa-miR-199a-3p	YP00204536
		hsa-miR-151a-3p	YP00204576
		hsa-miR-151a-5p	YP00204007
		hsa-miR-199a-5p	YP00204494
		hsa-miR-146a-5p	YP00204688
		hsa-miR-29b-3p	YP00204679
		hsa-miR-126-5p	YP00206010
Spike-ins	RNA extraction efficiency	UniSp2	YP00203950
		UniSp4	YP00203953
		UniSp5	YP00203955
	Reverse transcription efficiency	UniSp6	YP00203954
		cel-miR-39-3p	YP00203952
	Sample Signal	hsa-miR-142-3p	YP00204291
		hsa-miR-451a	YP02119305
		hsa-miR-23a-3p	YP00204772
		hsa-miR-30c-5p	YP00204783
		hsa-miR-103a-3p	YP00204063
		hsa-miR-191-5p	YP00204306

eTable2: Amplification efficiencies for Qiagen LNA miRCURY PCR assays calculated using LinRegPCR.

1. qPCR was run as described in Dunlop et al 2021 [4], then optical data exported with no baseline as real-time PCR data markup language (RMDL) v1.1 files and imported into LinRegPCR [6]. Individual amplification efficiencies were calculated for each well. The mean PCR efficiency for all nine primers was 1.86 +/- 0.013, meaning efficiencies varied by a maximum of three percent and we could compare results without the requirement for efficiency adjustments [7]. These results represent mean amplification efficiencies for each amplicon from all samples in all cohorts from this study. Although LNA-spiked PCR primers provide enhanced specificity which is especially beneficial when distinguishing between closely related miRNA sequences or detecting low-abundance targets, they are reported to have lower amplification efficiencies than DNA primers [8-10].

Qiagen LNA miRCURY PCR Assay	GeneGlobe ID	mean amplification efficiency	SD
hsa-miR-151a-3p	YP00204576	1.85	0.011
hsa-miR-151a-5p	YP00204007	1.85	0.015
hsa-miR-146a-5p	YP00204688	1.87	0.007
hsa-miR-4454	YP02114119	1.88	0.010
hsa-miR-10b-5p	YP00205637	1.87	0.008
hsa-miR-199a-3p	YP00204536	1.86	0.005
hsa-miR-199a-5p	YP00204494	1.86	0.007
hsa-miR-29b-3p	YP00204679	1.85	0.047
hsa-miR-126-5p	YP00206010	1.85	0.008
	Mean of all primers	1.86	0.013

eReferences

1. Banack SA, Dunlop RA, Cox PA. An miRNA fingerprint using neural-enriched extracellular vesicles from blood plasma: Towards a biomarker for amyotrophic lateral sclerosis/motor neuron disease. *Open Biol*. 2020;10(6). doi:10.1098/rsob.200116

2. Banack SA, Dunlop RA, Stommel EW, Mehta P, Cox PA. miRNA extracted from extracellular vesicles is a robust biomarker of amyotrophic lateral sclerosis. *J Neurol Sci*. 2022;442. doi:10.1016/j.jns.2022.120396

3. Banack SA, Dunlop RA, Mehta P, Mitsumoto H, Wood SP, Han M, Cox PA. A microRNA diagnostic biomarker for amyotrophic lateral sclerosis. *Brain Commun*. 2024;6(5):fcae268. doi:10.1093/braincomms/fcae268

4. Dunlop RA, Banack SA, Cox PA. A comparison of the efficiency of RNA extraction from extracellular vesicles using the Qiagen RNeasy MinElute versus Enzymax LLC RNA Tini Spin columns and qPCR of miRNA. *Biol Methods Protoc*. 2021;6(1):1-9. doi:10.1093/biomethods/bpab015

5. Dunlop RA, Banack SA, Cox PA. L1CAM immunocapture generates a unique extracellular vesicle population with a reproducible miRNA fingerprint. *RNA Biology. 2023;20*(1):140–148. doi.org/10.1080/15476286.2023.2198805

6. Untergasser A, Ruijter JM, Benes V, van den Hoff MJB. Web-based LinRegPCR: application for the visualization and analysis of (RT)-qPCR amplification and melting data. *BMC Bioinformatics*. 2021;22(1):398. doi:10.1186/s12859-021-04306-1

7. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods*. 2001;25(4):402-408. doi:10.1006/meth.2001.1262

8. Balcells I, Cirera S, Busk PK. Specific and sensitive quantitative RT-PCR of miRNAs with DNA primers. *BMC Biotechnol*. 2011;11:70. doi:10.1186/1472-6750-11-70

9. Levin JD, Fiala D, Samala MF, Kahn JD, Peterson RJ. Position-dependent effects of locked nucleic acid (LNA) on DNA sequencing and PCR primers. *Nucleic Acids Res*. 2006;34(20):e142. doi:10.1093/nar/gkl756

10. Brunet-Vega A, Pericay C, Quílez ME, Ramírez-Lázaro MJ, Calvet X, Lario S. Data on individual PCR efficiency values as quality control for circulating miRNAs. *Data Brief*. 2015;5:321-326. doi:10.1016/j.dib.2015.09.011