Circulating miR-181a-5p is a prognostic biomarker for amyotrophic lateral sclerosis

Iddo Magen[1], Anna Coenen-Stass[2], Nancy S. Yacovzada[1], Julian Grosskreutz[3,4], Ching-Hua Lu[5,6] Linda Greensmith[2], Andrea Malaspina[4]*, Pietro Fratta[2]* and Eran Hornstein[1]*

1 Department of Molecular Genetics, Weizmann Institute of Science, Rehovot, Israel

2 Department of Neuromuscular Diseases, UCL, Queen Square Institute of Neurology, London, UK

3 Hans Berger Department of Neurology, Jena University Hospital, Jena, Germany

4 Center for Healthy Aging, Jena University Hospital, Jena, Germany

5 Centre for Neuroscience and Trauma, Blizard Institute, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK.

6 School of Medicine, China Medical University, Taichung, Taiwan

* To whom correspondence should be addressed: a.malaspina@qmul.ac.uk, p.fratta@ucl.ac.uk or eran.hornstein@weizmann.ac.il

Abstract word count: 203

Main text word count: 2738 words (excluding references and figure legends).

No. of items (figures, tables): 4, 1

One Sentence Summary: higher plasma miR-181a-5p levels increase mortality risk in ALS patients.

Abstract (210 words)

Amyotrophic lateral sclerosis (ALS) is a ruthless neurodegenerative disease affecting the motor neuron system. Variability of disease progression has limited the effectiveness of ALS clinical trials, making novel tools for patient stratification at trial recruitment. Thus, better outcome measures are desperately needed in order to achieve therapeutic progress. Here, we investigate the potential of plasma cell-free microRNAs as biomarkers to predict ALS progression. We apply an unbiased high-throughput approach to define miRNA levels in a large cohort of ALS patients and controls. Crucially, we conduct our analysis also on longitudinal samples reflecting disease progression, and are able to integrate detailed clinical phenotyping in our analysis. We identify miR-181a-5p levels to be stable during disease course and demonstrate that high miR-181a-5p plasma levels predict shortened survival in ALS patients, with a 2.5 fold reduction in median survival for patients with high miR-181a-5p plasma levels. We replicated this finding in an independent validation cohort, where an eight-fold (×8) difference in miR-181a-5p levels between the two prognosis subgroups was robustly measurable by quantitative real time PCR. In conclusion, miR-181a-5p plasma levels can predict disease course in ALS suggesting it as a novel biomarker for patient stratification that might will greatly enhance the effectiveness of clinical trials.

Introduction

Amyotrophic lateral sclerosis (ALS) is a devastating neurodegenerative disease of the motor neuron system, for which no effective disease-modifying treatment exists. The disease is characterized by a significant variability in progression rates (*1, 2*) that imposes a challenge on patient stratification. Reliable biomarkers of disease progression would be valuable for ALS patient stratification prior to enrolment in clinical trials. Ideal biomarkers should remain stable during the course of disease and also be accessible, as in peripheral blood. To date, intensive research has identified only few peripheral blood ALS biomarkers (*3-5*), including cell-free neurofilaments (*6-8*), and pro-inflammatory cytokines (*9-11*).

microRNAs (miRNAs) are endogenous non-coding RNAs that are essential for motor neuron survival and are globally downregulated in post mortem ALS motor neurons (*12-14*). While circulating miRNA profiles have been previously characterized in ALS (*15-17*), the potential of miRNA biomarkers for ALS prognosis, and as readouts of disease progression has not yet been investigated. Here, we take an unbiased next generation sequencing (NGS) approach to comprehensively study plasma miRNAs in a large cohort of ALS cases and controls. We identified one miRNA, miR-181-5p, that is stable over time in individual patients. miR-181a-5p levels predict disease progression in an initial and second, validation cohort, demonstrating its use as a novel ALS biomarker.

Results

We used unbiased next generation sequencing to investigate, without an *a priori* bias, the comprehensive landscape of plasma miRNAs in ALS patients and controls, for whom in depth clinical information was available (Table 1).

First, we analyzed the differences between 115 ALS cases and 103 controls. Samples were prepared from total plasma RNA, as previously described (*18*), and profiled by small RNA next generation sequencing (RNA-seq). Linear miRNA quantification was achieved via unique 12-nucleotide molecular identifiers (UMIs). 98% of samples cleared our threshold of 50,000 total miRNA UMIs and were therefore included. miRNAs that were detected at ≥100 UMIs, on average across the 218 samples, were considered above noise level (a whole-cohort cut-off of ≥21,800 UMIs). Thus, of 1973 miRNAs aligned to the human genome (GRCh37/hg19), 216 miRNA species measured above noise level in at least one of the 218 samples, with an average of 184±5 miRNAs per sample above noise level (Supplementary table 1 describes the RNA-seq data).

We report 39 (/63) decreased (/increased) miRNAs in ALS samples relative to controls with adjusted p-values <0.05, Wald test, Figure S1A). However, none of the 102 miRNAs that were differentially expressed displayed satisfying diagnostic power. Specifically, three muscle-enriched miRNA were elevated in ALS plasma: miR-206 (2.2 fold, adjusted $p=6.6x10^{-6}$, Figure S1B), miR-133a-3p and miR-133b (~1.5 fold, adjusted p=0.01). The area under the curve (AUC) for receiver-operating characteristic (ROC) curve of miR-206 differential expression showed a modest predictive power for ALS diagnosis (AUC = 0.69, p<0.0001; Figure S1C), in accordance with previous studies (*19-21*). Beyond muscle-related miRNAs, miR-3168 was increased in ALS vs controls (2.3 fold increase; adjusted $p=6.6x10^{-6}$).

To experimentally assess miRNA measurement variability (noise), we analyzed a longitudinal sample cohort of 22 patients, for which blood samples were collected at four subsequent longitudinal time-points: t_1-t_4 during the course of 30 months (2.5 years). Measured variability between 22 individuals was low for 131/187 miRNAs (-log2 of t_4/t_1 standard error >-0.2), and these were considered stable, while 56/187 other miRNAs exhibited higher variability between patients and were excluded from further analysis (Figure 1A, *y-axis*). For example, miR-181a-5p levels exhibited reduced variability at each time point across individual patients, relative to miR-1-3p (Fig. 1B,C).

miRNAs could be further characterized by intra-personal changes over time. Intriguingly, four miRNAs (miR-423/484/92a/b) exhibited a robust increase in abundance from t_1 to t_4 (Figure 1A, *x-axis*; Z-score of log2 t_4/t_1 > 1.5 SD). miR-423/484/92a/b modestly increased early on (~1.15 fold t_2 or t_3 relative to t_1), and further increased ~1.7-fold upon the fourth sampling (t_4 / t_1, Figure S2A-C). The correlation of miR-423/484/92a/b levels to disease progression was observed also when time from enrolment (in months) was considered as a continuous variable (Figure S3A-I). Furthermore, miR-484/92a/b-3p levels increased significantly in a replication longitudinal cohort (N=22-24 patients, one-tailed t-test: miR-484, p=0.01; miR-92a-3p, p=0.006; miR-92b-3p, p=0.04; Figure S3K-M), whereas miR-423-5p did not replicate the upward trend in the second longitudinal cohort (one-tailed p-value=0.16, Figure S3J).

The relative D50 (rD50), is a derivative of ALSFRS-R, that reveals the disease covered by individual patients (*22*). E.g., rD50 of 0.0 and 0.5 signify ALS onset and the time-point of halved functionality, respectively (*22*). miR-484/92a/b levels measured at longitudinal blood samplings from the same patients correlated with rD50 at the time of sampling (Figure S4) and

may be considered as candidate molecular biomarkers of the individual disease covered and thus of ALS progression.

We next explored potential blood-borne miRNAs for prognosis biomarkers. An ideal prognosis biomarker should be preferentially stable during the course of disease. We identified 127 miRNA candidate biomarkers, whose plasma levels were relatively stable during disease progression (Figure 1A). We then correlated miRNA expression at enrolment to survival duration, calculated as lifespan from either symptom onset or from first phlebotomy at study enrolment (on average 30.0±6 months after symptom onset). The levels of one miRNA, miR-181a-5p exhibited significant correlation to survival length (Spearman correlation coefficient -0.21, p=0.02, Figure 2A, S5A). Because miR-181a-5p levels remain stable in time (Figure 1A,B), it is unlikely that measurements are confounded by sampling at different disease stages. The stability of miR-181a-5p during disease progression is also indirectly implied by comparable miR-181a-5p levels between patients sampled at different (early, progressive and late) disease phases ($0 \leq rD50 < 0.25$; $0.25 \leq rD50 < 0.5$; $rD50 \geq 0.5$, respectively; ANOVA p=0.35, Figure S6A) and by lack of correlation between miR-181a-5p levels and rD50 at enrolment (Fig. S6B).

We further subdivided patients in quartiles by the expression level of individual miRNA genes, and investigated how the low quartile (30 patients) compared to the rest of the cases (85 patients). Differences in plasma miR-181a-5p levels were able to predict survival length better than any other miRNA. Kaplan Meier curves revealed clear separation of survival between the subgroups, based on plasma miR-181a-5p levels at enrolment (log-rank test chi-square=8.6, p=0.003, Figure 2B). The median patient survival in the lower miR-181a-5p quartile was 36 months, compared to 13 months at the higher three quartiles. Thus, plasma miR-181a levels predict a substantial median survival difference of 23 months that is equivalent to a 280%

increase in survival length for patients with lower plasma miR-181a-5p levels. Comparable results were gained, when plasma miR-181a-5p levels at enrolment were stratified against survival from reported first symptoms (disease onset, log-rank chi-square=9.4, p=0.002, Figure 2C).

Noteworthy, patient individual disease covered at enrolment (rD50) was comparable between the low quartile of miR-181a-5p and the rest of the cases (Figure S6C). Furthermore, enrolment progression rate, enrolment ALSFRS-R scores, and chronological age at onset were not correlated with miR-181a-5p levels, and these phenotypic features did not differ between the low and high miR-181a-5p quantiles, except for a significant but modest increase in enrolment progression rate (Figure S7A-F). These findings rule out the possibility that survival differences between miR-181a-5p expression bins result from factors other than miR-181a-5p expression levels, such as the clinical status of the patients when they are sampled.

The monthly mortality hazard ratio (HR) was calculated for different plasma miR-181a-5p quantiles and clinical covariates, by multivariate Cox regression analysis. This analysis allows to calculate an independent hazard ratio for each covariate while holding the other covariates constant. High miR-181a-5p levels (top three quartiles) exhibited a hazard ratio of 2.35 (95% CI 1.4-3.9), relative to the lowest quartile when calculated from enrolment (Wald test $p<0.001$, Figure 3A), and a hazard ratio of 2.4 when calculated from disease onset (95% CI, 1.5-4.0, $p<0.01$, Figure 3B). Enrolment progression rate, i.e. the rate of change in ALSFRS from symptom onset to enrolment, was also associated with a high hazard ratio of 2.16 for survival time from enrolment (95% CI 1.6-2.9, $p<0.001$, Figure 3A), as previously reported (*23*). Taken together, higher enrolment miR-181a plasma levels associate with a considerably poor prognosis in ALS, an effect which is not confounded by other phenotypic properties.

In order to validate our findings, we performed a replication study by measuring miR-181a-5p levels in plasma of an independent cohort of 56 patients (Table 1) by an orthogonal qPCR methodology. The replication study confirmed the existence of a prognostic gap of 180%, from 14 to 25 months, between patients with high or low plasma miR-181a-5p levels, respectively (Log rank test: chi square=3.45, one-tailed p-value=0.03, Figure 4A). The survival difference was reproduced in the replication cohort, even when calculated from first reported symptoms (230%, 31 months from disease onset to 71 months, chi square=3.0, one-tailed p-value=0.04, Figure 4B). Importantly, the average miR-181a-5p levels varied by 8-fold between the plasma of patients with better prognosis, relative to those with poor prognosis (Figure 4C), suggesting a good dynamic range for biomarker quantification. Our data from two different cohorts of ALS patients establish miR-181a-5p as a new cell-free prognostic biomarker for ALS.

In summary, we report the results of the most elaborated small RNA-seq study in ALS undertaken to date. Our work confirms the increase in muscle-enriched miR-206/133, which plausibly reflects myotube content discharge from the muscular atrophy in ALS. We further show that miRNAs appear to mostly be unchanged longitudinally during disease, with the exception of miR-423/484/92a/b, that increase over 30 months of disease. This feature could be used to monitor disease progression / individual disease covered and warrants further exploration as pharmacodynamic biomarkers of response to therapy. Crucially, we find that miR-181a-5p levels predict shortened survival in two separate ALS cohorts. Intriguingly, the median survival of patients with low miR-181a-5p plasma levels was more than double that of patients with high miR-181a-5p. An eight-fold difference in miR-181a-5p levels between the two prognosis subgroups, is robustly measurable. Although the tissue source of miR-181a-5p is unknown, it

emerges as a prognostic ALS biomarker that is able to improve patient stratification in clinical trials.

Materials and Method

Standard protocol approvals, registrations, and patient consents

All participants provided written consent (or gave verbal permission for a carer to sign on their behalf) to be enrolled in the ALS biomarkers study and ethical approval was obtained from the East London and the City Research Ethics Committee 1 (09/H0703/27).

Study design

Based on power analysis, we found that about 100 control and 100 ALS samples are required to obtain an ROC of 0.7 with a power of 99% and a p-value of 0.01. For survival analysis, 120 ALS patients were needed to obtain a hazard ratio of 3-fold in short survivors compared with long survivors, , with the same power and p-value. We determined the sample size based on these calculations. Because sample processing was done in different batches, samples were randomly allocated to the batches and within each batch, the number of control and ALS samples was balanced in order to reduce batch-associated bias.

Participants and sampling

This study included a first cohort with 115 patients with ALS and 103 healthy controls from the ALS biomarker study. ALS patients were diagnosed according to standard criteria by experienced ALS neurologists (*24*). Healthy controls were typically spouses or relatives of patients. Serial plasma samples and clinical information were obtained, on average, every 2 to 4 months from 22 of the 115 patients with ALS. No selection criteria were applied to individuals with ALS sampled longitudinally, other than their willingness to donate further samples. Symptom onset was defined as first patient-reported weakness. Disease severity was assessed with the revised ALS Functional Rating Scale (ALSFRS-R)(*25*), and progression rate at

enrolment (i.e. first blood draw) was calculated as follows: (48 - enrolment ALSFRS-R)/time (in months) from symptom onset to enrolment. Progression was also modeled using the D50 model (*22, 26*). Use of Riluzole (or not) at the time of sampling was recorded. In samples from cohort I (115 ALS patients, 103 healthy controls), miRNAs were profiled with NGS. A second cohort of 56 ALS patients, was used to validate the results of the first cohort by an orthogonal method of quantitative real time PCR. Plasma samples were stored in -80^{0}C until RNA extraction and subsequent small RNA next generation sequencing.

Small RNA next generation sequencing

Total RNA was extracted from plasma using the miRNeasy micro kit (Qiagen, Hilden, Germany) and quantified with Qubit fluorometer using RNA broad range (BR) assay kit (Thermo Fisher Scientific, Waltham, MA). For small RNA next generation sequencing (NGS), libraries were prepared from 7.5 ng of total RNA using the QIAseq™ miRNA Library Kit and QIAseq miRNA NGS 48 Index IL (Qiagen), by an experimenter who was blinded to the identity of samples. Following 3' and 5' adapter ligation, small RNA was reverse transcribed, using unique molecular identifier (UMI), primers of random 12-nucleotide sequences. This way, precise linear quantification miRNA is achieved, overcoming potential PCR-induced biases (*18*). cDNA libraries were amplified by PCR for 22 cycles, with a 3' primer that includes a 6-nucleotide unique index. Following size selection and cleaning of libraries with magnetic beads, quality control was performed by measuring library concentration with Qubit fluorometer using dsDNA high sensitivity (HS) assay kit (Thermo Fisher Scientific, Waltham, MA) and confirming library size with Tapestation D1000 (Agilent). Libraries with different indices were multiplexed and sequenced on a single NextSeq 500/550 v2 flow cell (Illumina), with 75bp single read and 6bp index read. Fastq files were demultiplexed using the User-friendly Transcriptome Analysis

Pipeline (UTAP) developed at the Weizmann Institute (*27*). Human exact mature miRNA sequence, as defined by miRBase (*28*), were mapped using Geneglobe (Qiagen).

Quantitative real time PCR (qPCR)

TaqMan Advanced MicroRNA cDNA Synthesis Kit (Applied Biosystems) was used for cDNA reverse transcription (10 ng input). Universal PCR Master Mix without AmpErase UNG on white reaction plates (MicroAmp EnduraPlate Optical 96-well, Thermo Fisher) and run on a StepOnePlus machine (Applied Biosystems). TaqMan qPCR was performed using Advanced MicroRNA Assays (reaction volume 10 μL) with the following probes: hsa-miR-181a-5p (Assay ID: 477857_mir); hsa-miR-423-5p (Assay ID: 478090_mir); hsa-miR-484 (Assay ID: 478308_mir); hsa-miR-92a-3p (Assay ID: 477827_mir); hsa-miR-92b-3p (Assay ID: 477823_mir); hsa-miR-140-3p (Assay ID: 477908_mir) and hsa-miR-185-5p (Assay ID: 477939_mir) were selected as normalizers based on stable expression in the first cohort: (1) basemean expression between 500-3,000; (2) coefficient of variation ≤ 0.35 (*29*) and (3) minimal, insignificant, changes in patient survival between different expression bins of the miRNA. Relative miR-181a-5p quantity, normalized to the average expression of hsa-miR-140-3p and hsa-miR-185-5p were binned to lowest 25% and highest 75%, for Kaplan-Meier survival analysis. Relative miR-423/484/92a/92b quantities, normalized to same normalizers was compared between enrolment sample (t_1) and corresponding follow-up sample (t_2) by one-tailed paired t-test.

Statistical analysis

Plasma samples with ≥50,000 total miRNA UMIs were included. miRNA with average abundance of ≥100 UMIs per sample, across all samples, were considered above noise levels. Thus, the whole-cohort UMI cut-off was 100 ×218 = 21,800 UMIs). miRNA NGS data was analyzed via DESeq2 package in R Project for Statistical Computing (*30, 31*), under the assumption that miRNA counts followed negative binomial distribution and data were corrected for library preparation batch in order to reduce its potential bias. For longitudinal multiple time points analysis, correction was performed per subject. Ratio of normalized ALS counts to the normalized control counts presented after logarithmic transformation on base 2. *P* values were calculated by Wald test (*31, 32*) and adjusted for multiple testing according to Benjamini and Hochberg (*33*). Log rank Mantel-Cox test was used for Kaplan-Meier survival estimators and a fixed date was used to censor data for survival analysis. Cox regression was used to calculate mortality hazard ratio, with miR-181a-5p levels and phenotypic features covariates. Diagnosis by miR-206 levels performed with receiver operating characteristic (ROC) curves on voom transformation of gene expression data package in R (*34*). For longitudinal miRNA expression by qPCR, one-tailed paired t-test was used. Outliers were detected by Grubbs test (*35*) and excluded from analysis. Graphs were generated with GraphPad Prism 5.

Acknowledgements

We thank Vittoria Lombardi (UCL) for technical assistance, and Iddo Ben-Dov from Hadassah Hebrew University Medical Center, Jerusalem, Israel, for advice on statistics. We acknowledge patients for their contribution and all the ALS biomarkers study co-workers and their contribution to the biobanking project which has made this study possible (REC 09/H0703/27).

Funding

This research was supported by the following grants: Motor Neuron Disease Association (MNDA) 839-791; ISF Legacy 828/17; TEVA NNE postdoc Fellowship to IM 117941; PF is supported by an MRC/MND LEW Fellowship and by the NIHR UCLH BRC.

Competing interests

The authors state that they have no competing interests.

References

1. T. L. Munsat, P. L. Andres, L. Finison, T. Conlon, L. Thibodeau, The natural history of motoneuron loss in amyotrophic lateral sclerosis. *Neurology* **38**, 409-413 (1988).
2. J. M. Ravits, A. R. La Spada, ALS motor phenotype heterogeneity, focality, and spread: deconstructing motor neuron degeneration. *Neurology* **73**, 805-811 (2009).
3. M. Benatar, K. Boylan, A. Jeromin, S. B. Rutkove, J. Berry, N. Atassi, L. Bruijn, ALS biomarkers for therapy development: State of the field and future directions. *Muscle Nerve* **53**, 169-182 (2016).
4. R. H. Chipika, E. Finegan, S. Li Hi Shing, O. Hardiman, P. Bede, Tracking a Fast-Moving Disease: Longitudinal Markers, Monitoring, and Clinical Trial Endpoints in ALS. *Front Neurol* **10**, 229 (2019).
5. N. S. Verber, S. R. Shepheard, M. Sassani, H. E. McDonough, S. A. Moore, J. J. P. Alix, I. D. Wilkinson, T. M. Jenkins, P. J. Shaw, Biomarkers in Motor Neuron Disease: A State of the Art Review. *Front Neurol* **10**, 291 (2019).
6. D. Rossi, P. Volanti, L. Brambilla, T. Colletti, R. Spataro, V. La Bella, CSF neurofilament proteins as diagnostic and prognostic biomarkers for amyotrophic lateral sclerosis. *J Neurol* **265**, 510-521 (2018).
7. M. Goncalves, M. De Carvalho, C. Peixoto, P. Alves, C. Barreto, A. Oliva, S. Pinto, A. Laborinho-Pronto, M. Gromicho, J. Costa, Phosphoneurofilament heavy chain and vascular endothelial growth factor as cerebrospinal fluid biomarkers for ALS. *Amyotroph Lateral Scler Frontotemporal Degener* **18**, 134-136 (2017).
8. C. H. Lu, C. Macdonald-Wallis, E. Gray, N. Pearce, A. Petzold, N. Norgren, G. Giovannoni, P. Fratta, K. Sidle, M. Fish, R. Orrell, R. Howard, K. Talbot, L. Greensmith, J. Kuhle, M. R. Turner, A. Malaspina, Neurofilament light chain: A prognostic biomarker in amyotrophic lateral sclerosis. *Neurology* **84**, 2247-2257 (2015).
9. M. T. Gonzalez-Garza, H. R. Martinez, D. E. Cruz-Vega, M. Hernandez-Torre, J. E. Moreno-Cuevas, Adipsin, MIP-1b, and IL-8 as CSF Biomarker Panels for ALS Diagnosis. *Dis Markers* **2018**, 3023826 (2018).
10. L. G. R. Prado, N. P. Rocha, L. C. de Souza, I. C. S. Bicalho, R. S. Gomez, M. Vidigal-Lopes, N. F. T. Braz, E. L. M. Vieira, A. L. Teixeira, Longitudinal assessment of clinical and inflammatory markers in patients with amyotrophic lateral sclerosis. *J Neurol Sci* **394**, 69-74 (2018).
11. C. H. Lu, K. Allen, F. Oei, E. Leoni, J. Kuhle, T. Tree, P. Fratta, N. Sharma, K. Sidle, R. Howard, R. Orrell, M. Fish, L. Greensmith, N. Pearce, V. Gallo, A. Malaspina, Systemic inflammatory response and neuromuscular involvement in amyotrophic lateral sclerosis. *Neurol Neuroimmunol Neuroinflamm* **3**, e244 (2016).
12. S. Haramati, E. Chapnik, Y. Sztainberg, R. Eilam, R. Zwang, N. Gershoni, E. McGlinn, P. W. Heiser, A. M. Wills, I. Wirguin, L. L. Rubin, H. Misawa, C. J. Tabin, R. Brown, Jr., A. Chen, E. Hornstein, miRNA malfunction causes spinal motor neuron disease. *Proc Natl Acad Sci U S A* **107**, 13111-13116 (2010).
13. A. Emde, C. Eitan, L. L. Liou, R. T. Libby, N. Rivkin, I. Magen, I. Reichenstein, H. Oppenheim, R. Eilam, A. Silvestroni, B. Alajajian, I. Z. Ben-Dov, J. Aebischer, A. Savidor, Y. Levin, R. Sons, S. M. Hammond, J. M. Ravits, T. Moller, E. Hornstein, Dysregulated miRNA biogenesis downstream of cellular stress and ALS-causing mutations: a new mechanism for ALS. *EMBO J* **34**, 2633-2651 (2015).
14. C. Eitan, E. Hornstein, Vulnerability of microRNA biogenesis in FTD-ALS. *Brain Res*, (2016).
15. I. Takahashi, Y. Hama, M. Matsushima, M. Hirotani, T. Kano, H. Hohzen, I. Yabe, J. Utsumi, H. Sasaki, Identification of plasma microRNAs as a biomarker of sporadic Amyotrophic Lateral Sclerosis. *Mol Brain* **8**, 67 (2015).

16. H. M. de Andrade, M. de Albuquerque, S. H. Avansini, S. R. C. de, D. B. Dogini, A. Nucci, B. Carvalho, I. Lopes-Cendes, M. C. Franca, Jr., MicroRNAs-424 and 206 are potential prognostic markers in spinal onset amyotrophic lateral sclerosis. *J Neurol Sci* **368**, 19-24 (2016).
17. K. S. Sheinerman, J. B. Toledo, V. G. Tsivinsky, D. Irwin, M. Grossman, D. Weintraub, H. I. Hurtig, A. Chen-Plotkin, D. A. Wolk, L. F. McCluskey, L. B. Elman, J. Q. Trojanowski, S. R. Umansky, Circulating brain-enriched microRNAs as novel biomarkers for detection and differentiation of neurodegenerative diseases. *Alzheimers Res Ther* **9**, 89 (2017).
18. A. M. L. Coenen-Stass, I. Magen, T. Brooks, I. Z. Ben-Dov, L. Greensmith, E. Hornstein, P. Fratta, Evaluation of methodologies for microRNA biomarker detection by next generation sequencing. *RNA Biol* **15**, 1133-1145 (2018).
19. E. Tasca, V. Pegoraro, A. Merico, C. Angelini, Circulating microRNAs as biomarkers of muscle differentiation and atrophy in ALS. *Clin Neuropathol* **35**, 22-30 (2016).
20. J. M. Toivonen, R. Manzano, S. Olivan, P. Zaragoza, A. Garcia-Redondo, R. Osta, MicroRNA-206: a potential circulating biomarker candidate for amyotrophic lateral sclerosis. *PLoS One* **9**, e89065 (2014).
21. R. Waller, E. F. Goodall, M. Milo, J. Cooper-Knock, M. Da Costa, E. Hobson, M. Kazoka, H. Wollff, P. R. Heath, P. J. Shaw, J. Kirby, Serum miRNAs miR-206, 143-3p and 374b-5p as potential biomarkers for amyotrophic lateral sclerosis (ALS). *Neurobiol Aging* **55**, 123-131 (2017).
22. T. Prell, B. Stubendorff, T. T. Le, N. Gaur, V. Tadic, A. Rodiger, O. W. Witte, J. Grosskreutz, Reaction to Endoplasmic Reticulum Stress via ATF6 in Amyotrophic Lateral Sclerosis Deteriorates With Aging. *Front Aging Neurosci* **11**, 5 (2019).
23. M. Elamin, P. Bede, A. Montuschi, N. Pender, A. Chio, O. Hardiman, Predicting prognosis in amyotrophic lateral sclerosis: a simple algorithm. *J Neurol* **262**, 1447-1454 (2015).
24. B. R. Brooks, R. G. Miller, M. Swash, T. L. Munsat, D. World Federation of Neurology Research Group on Motor Neuron, El Escorial revisited: revised criteria for the diagnosis of amyotrophic lateral sclerosis. *Amyotroph Lateral Scler Other Motor Neuron Disord* **1**, 293-299 (2000).
25. J. M. Cedarbaum, N. Stambler, E. Malta, C. Fuller, D. Hilt, B. Thurmond, A. Nakanishi, The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. BDNF ALS Study Group (Phase III). *J Neurol Sci* **169**, 13-21 (1999).
26. K. Poesen, M. De Schaepdryver, B. Stubendorff, B. Gille, P. Muckova, S. Wendler, T. Prell, T. M. Ringer, H. Rhode, O. Stevens, K. G. Claeys, G. Couwelier, A. D'Hondt, N. Lamaire, P. Tilkin, D. Van Reijen, S. Gourmaud, N. Fedtke, B. Heiling, M. Rumpel, A. Rodiger, A. Gunkel, O. W. Witte, C. Paquet, R. Vandenberghe, J. Grosskreutz, P. Van Damme, Neurofilament markers for ALS correlate with extent of upper and lower motor neuron disease. *Neurology* **88**, 2302-2309 (2017).
27. R. Kohen, J. Barlev, G. Hornung, G. Stelzer, E. Feldmesser, K. Kogan, M. Safran, D. Leshkowitz, UTAP: User-friendly Transcriptome Analysis Pipeline. *BMC bioinformatics* **24**, 154 (2019).
28. A. Kozomara, S. Griffiths-Jones, miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* **42**, D68-73 (2014).
29. Z. Zhou, P. Cong, Y. Tian, Y. Zhu, Using RNA-seq data to select reference genes for normalizing gene expression in apple roots. *PLoS One* **12**, e0185288 (2017).
30. R Core Team, R: A Language and Environment for Statistical Computing. Vienna, Austria. *R Foundation for Statistical Computing* (2015).
31. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550 (2014).
32. S. Anders, W. Huber, Differential expression analysis for sequence count data. *Genome Biol* **11**, R106 (2010).
33. Y. Benjamini, D. Drai, G. Elmer, N. Kafkafi, I. Golani, Controlling the false discovery rate in behavior genetics research. *Behav Brain Res* **125**, 279-284 (2001).

34. C. W. Law, Y. Chen, W. Shi, G. K. Smyth, voom: Precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol* **15**, R29 (2014).

35. F. E. Grubbs, Procedures for Detecting Outlying Observations in Samples. *Technometrics* **11**, 1-21 (1969).

Magen et al Table 1

	ALS Cohort I	ALS Cohort II	Controls
Number of subjects (% males)	115 (64.3%)	56 (48.2%)	103 (27.2%)
Age at enrolment	64±1 yr.	65±2 yr.	52±2 yr.
Age of onset (1st reported symptoms)	61±1 yr.	62±2 yr.	N/A
Disease duration at enrolment	30±3 m.	35±6 m.	N/A
ALSFRS-R at enrolment	36±1	37±1	N/A
Progression rate at enrolment	0.8±0.1	0.9±0.2	N/A
El-Escorial (Definite/probable/lab-supported/possible/suspected)	34/51/15/15/0	3/14/21/15/3	N/A
Bulbar onset / total	38/115	17/56	N/A
Riluzole treated / total	80/115	34/56	N/A
ALS Family history / total	7/115	unknown	N/A
C9ORF72 genetics / total	8/115	2/56	N/A

Table 1. Summary of demographic and clinical characteristics of ALS Cohorts I and II and control samples. ALSFRS-R: ALS functional rating scale. Demographic data related to supplementary Figure 1, but not for main text, recognizes that male ratio and age of first phelbotomy was significantly different between ALS cohort I and controls (proportion test: $p<0.0001$; t-test: $p<0.001$, respectively).

Magen et al Figure 1

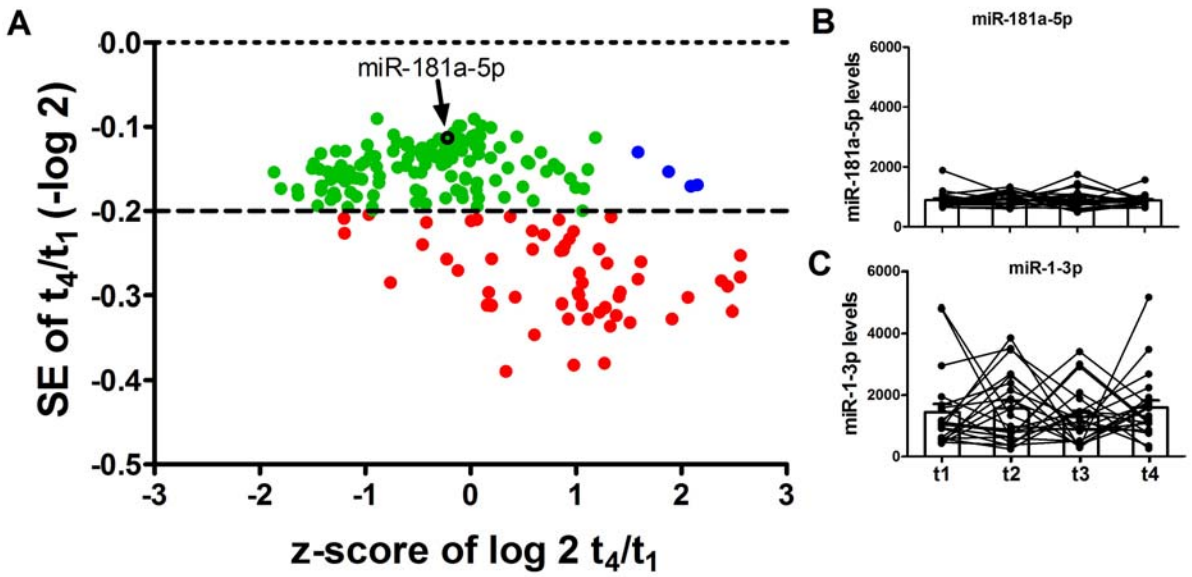


Figure 1. Assessment of miRNAs signal and noise in longitudinal samples from Patients with ALS. (A) The x-axis denotes the normalized change in miRNA levels between first and last measurements (as log 2-transformed t_4/t_1 ratios), relative to the average change of all 187 sequenced miRNAs. The y-axis denotes the variability in measurements, per-miRNA for 187 sequenced species between the 22 individuals (- log 2-transformed values ddof the standard error of t_4/t_1 ratios. Green, red features: are above or below the threshold of stability, respectively (-0.2). Blue features: stable miRNAs (miR-423/484/92a/b) that are increasing progressively with disease in individual pateints > 1.5 SD from the average miRNA. **(B-C)** Spaghetti plots of individual trajectories over four longitudinal sample collections for **(B)** miR-181a-5p and **(C)** miR-1-3p.

Magen et al Figure 2

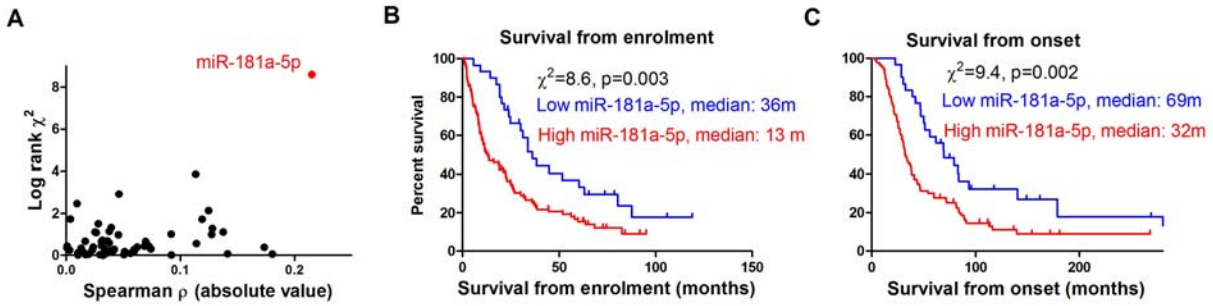


Figure 2. Prognostic value of miR-181a-5p in predicting ALS patient survival. (A) Survival predictive value of individual miRNA levels in plasma, binned by lower quantile vs. higher 3 quantiles (logrank test, *y-axis*), for miRNAs depicted in Figure 1 as 'stable'. Data plotted against miRNA level correlation to survival length (absolute value of Spearman correlation-coefficient, *x-axis*). Cumulative survival (Kaplan-Meier) curves from **(B)** enrolment, or from **(C)** reported first symptoms for patients, subdivided by miR-181a-5p levels, at low quartile (30 patients) compared to the rest of the cases (85 patients).

Magen et al Figure 3

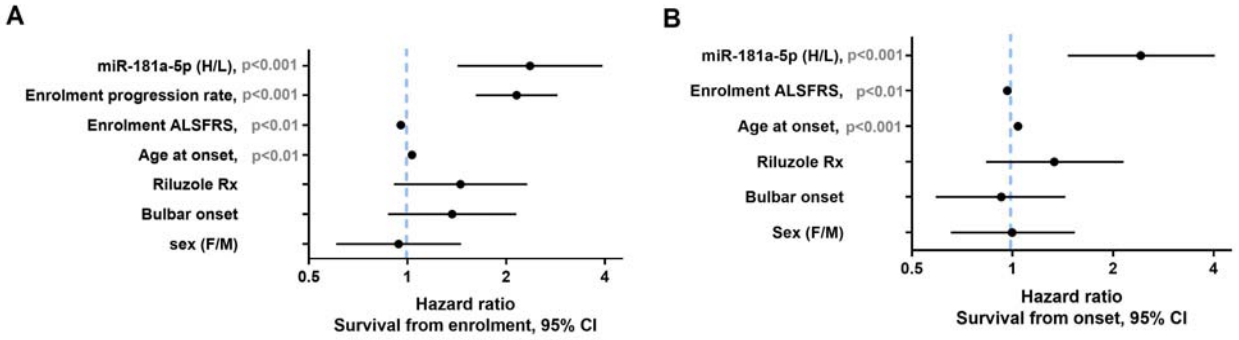


Figure 3. Mortality hazard ratio for miR-181a-5p. Mortality hazard ratio calculated by survival length from enrolment **(A)**, or from disease onset **(B).** Cox-regression analyses of covariates: miR-181a-5p levels, enrolment progression rate, enrolment ALSFRS-R score, age at onset, Riluzole-treatment, Bulbar onset, sex (Female/Male). Light blue dashed-line denotes unchanged mortality. Mean ± 95% CIs.

Magen et al Figure 4

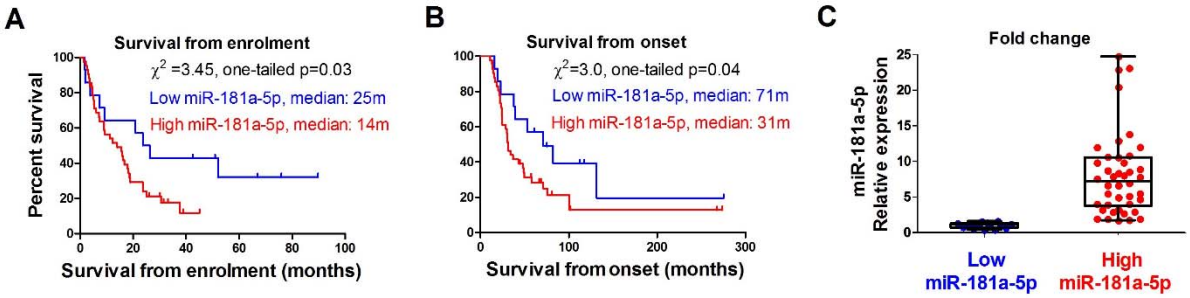


Figure 4. Prognostic value of miR-181a-5p is replicated in an independent cohort and by an orthogonal measuring technique. Kaplan-Meier curves from **(A)** enrolment, or **(B)** from reported first symptoms for a replication cohort, subdivided by miR-181a-5p levels, at low (14 patients) vs. high (42 patients) quantiles. (**c**) Scatter plot of plasma miR-181a-5p levels at low vs. high quantiles.

Supplementary figures and tables

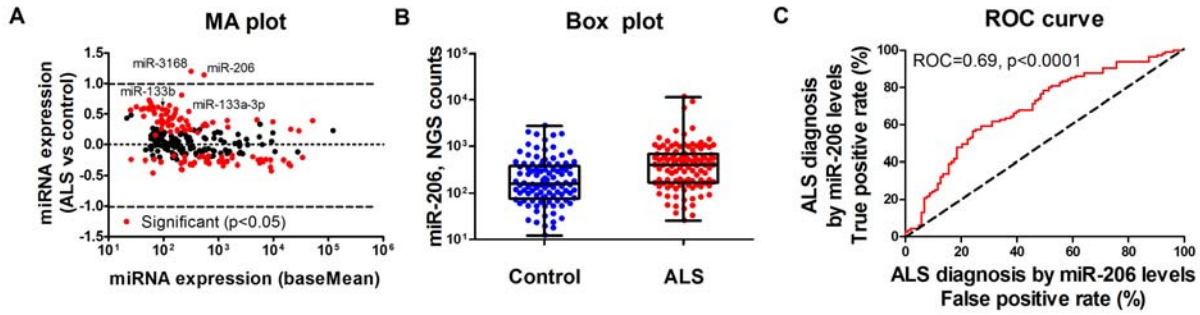


Figure S1. Differential miRNA expression in ALS plasma. **(A)** MA plot of differential miRNA expression in ALS (n=115 plasma samples) and healthy controls (n=103 plasma samples). Log-2 transformed fold-change (y-axis), against mean miRNA abundance (x-axis). Red -significantly changed miRNAs (adjusted p-value ≤0.05). 39 miRNAs decreased and 63 miRNAs increased in ALS samples, relative to controls. **(B)** Scatter plot of miR-206 counts in NGS data, normalized and transformed with the DESeq2 algorithm, for control and ALS. **(C)** Receiver operating characteristic (ROC) curve of miR-206 levels capacity to distinguish controls from ALS. True positive rates (sensitivity) as a function of false positive rates (1-specificity) for different cut-off values. P-values are calculated given null hypothesis of area under the curve (AUC) =0.5.

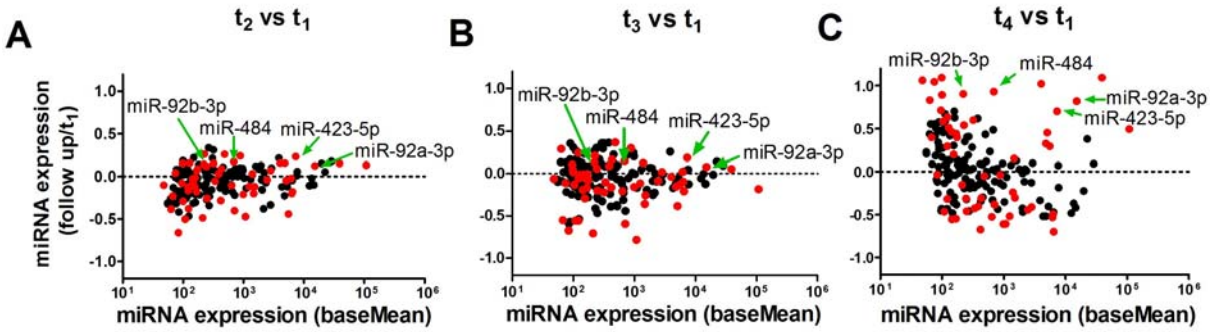


Figure S2. (A-C) MA plots of differential miRNA expression upon repeated sampling relative to the first phlebotomy. Red features denote miRNAs with statistically significant change in levels.

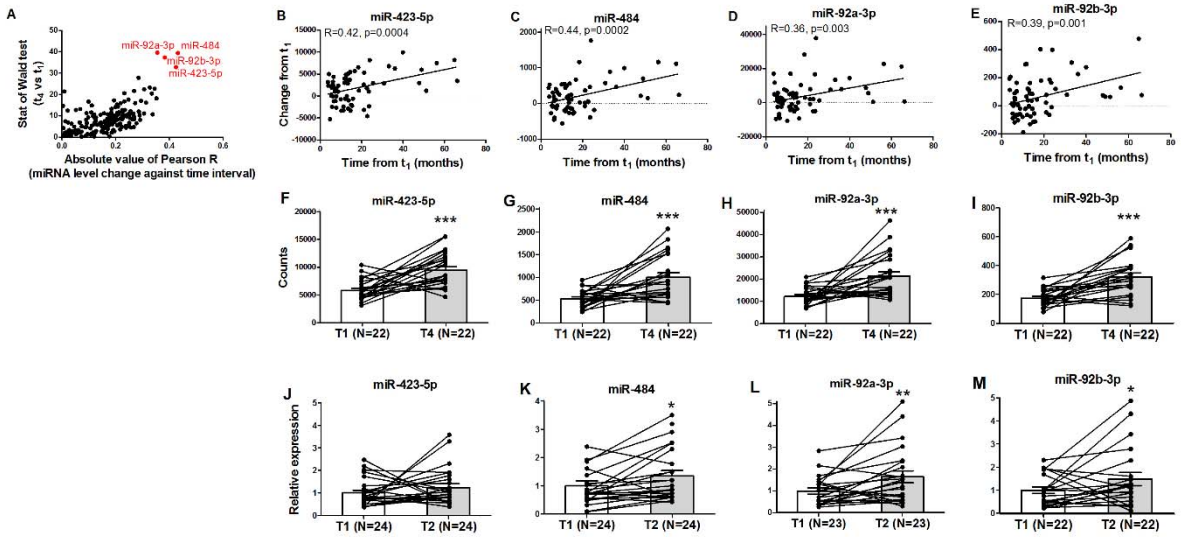


Figure S3. Longitudinal changes in miRNAs in ALS. (A) Differential expression between t_4 and t_1 of individual miRNA levels in plasma, (Wald test statistics, *y-axis*), plotted against correlation between change in miRNA level from t_1 and time interval from t_1 (absolute value of Pearson correlation-coefficient, *x-axis*). Temporal changes in the levels of **(B)** miR-423-5p, **(C)** miR-484, **(H)** miR-92a-3p, or **(E)** miR-92b-3p and **(F-I)** and corresponding spaghetti plots of individual trajectories (t_1-t_4). Time intervals: t_1-t_2 6.3±0.3 m.; t_1-t_3 13.0±0.3 m.; t_1=t_4 32.7±3 m. Disease duration: t_1 28.8±3 m.; t_4 61.5±3 m. Replication study of **(J)** miR-423-5p **(K)** miR-484 **(L)** miR-92a-3p or **(M)** miR-92b-3p levels by qPCR at second/first phlebotomy (t_1-t_2) in individual patients. Mean ± SEM. *p<0.05, **p<0.01, ***p<0.001, one-tailed paired t-test or Wald test.

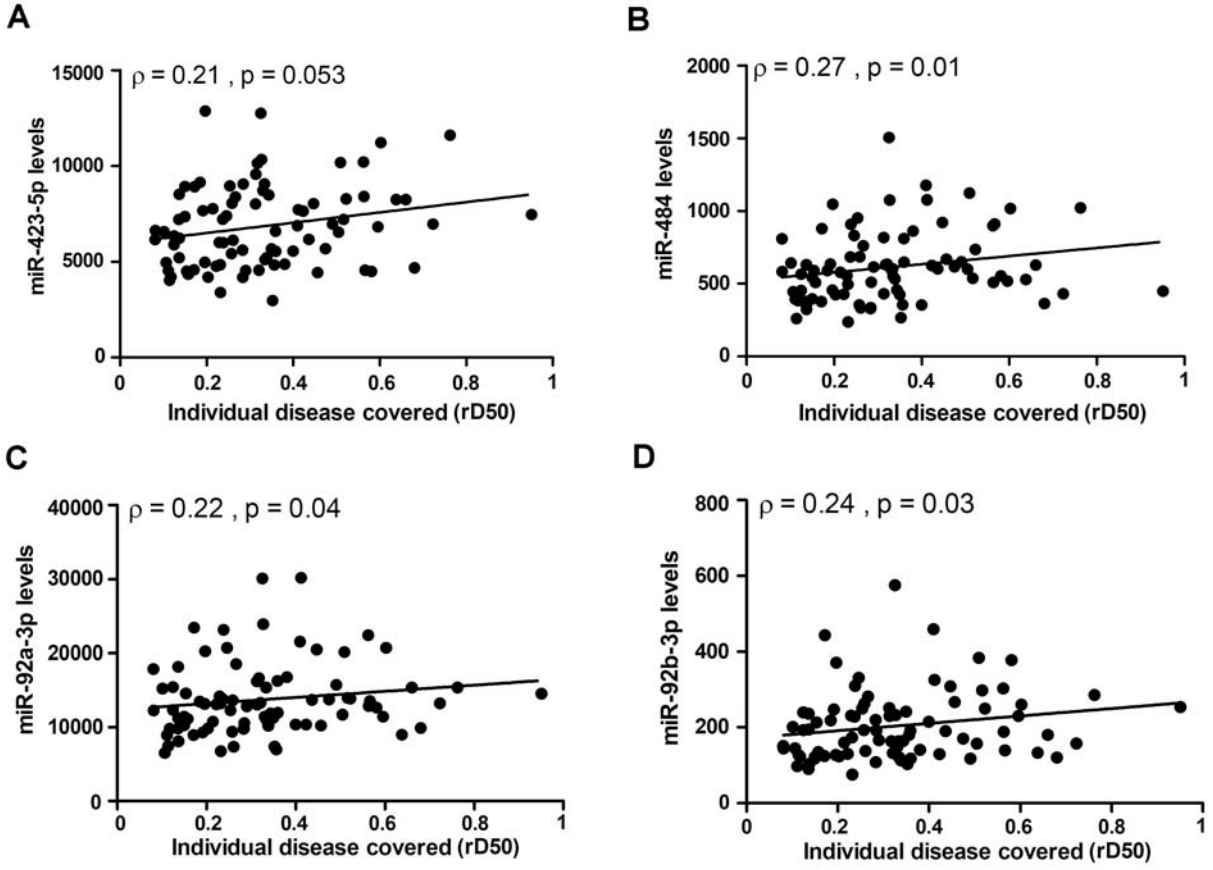


Figure S4. Longitudinally changed miRNAs increase with disease progression. Correlation between the relative disease covered (rD50) in longitudinal plasma collections (X-axis) and levels of **(A)** miR-423-5p, **(B)** miR-484, **(C)** miR-92a-3p, and **(D)** miR-92b-3p.

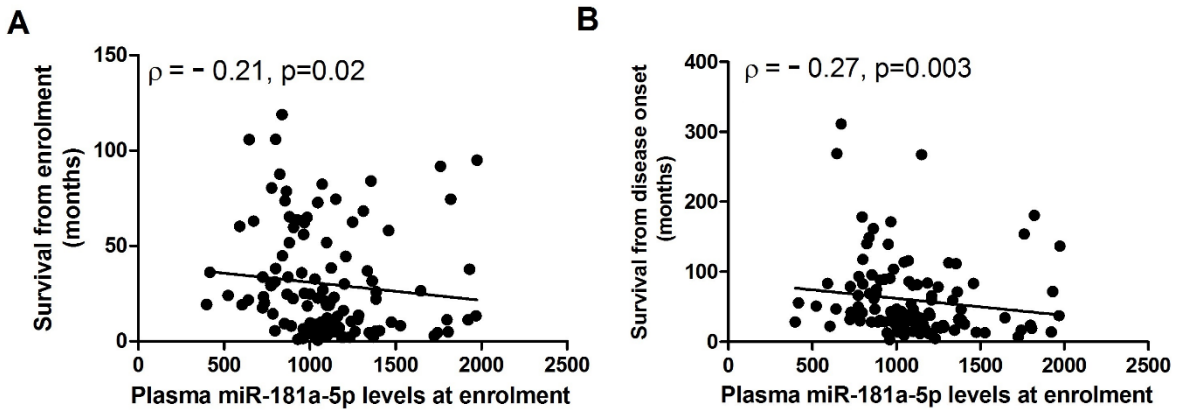


Figure S5. Plots depicting inverse correlation between miR-181a-5p and survival from first phlebotomy (A), or from disease onset (B). When the three samples taken from subjects whose survival from onset was above 260 months, Spearman correlation slightly decreases (from -0.27 to -0.25) and the p-values marginally increases (from 0.003 to 0.008), suggesting the overall data and interpretation remain unchanged.

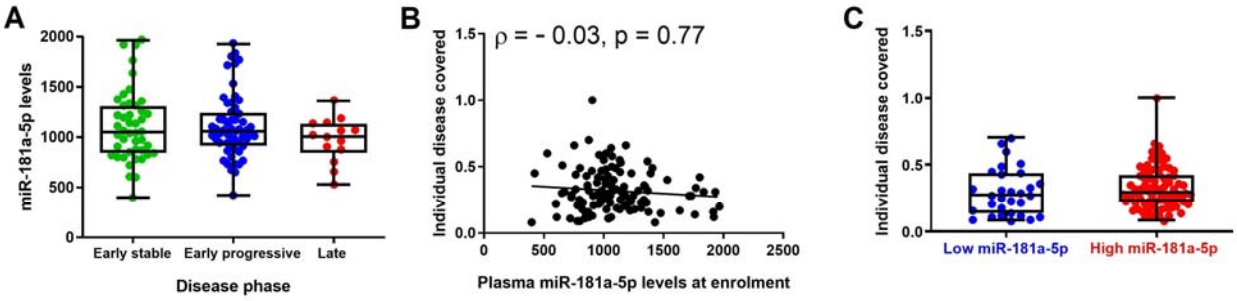


Figure S6. miR-181a-5p levels are not associated with disease phase. **(A)** miR-181a-5p in different phases of disease defined by rD50 values. One-way ANOVA: F (2,112) = 1.06, p=0.35. **(B)** correlation of miR-181a-5p levels with rD50 **(C)** rD50 differences between low and high miR-181a-5p expression bins.

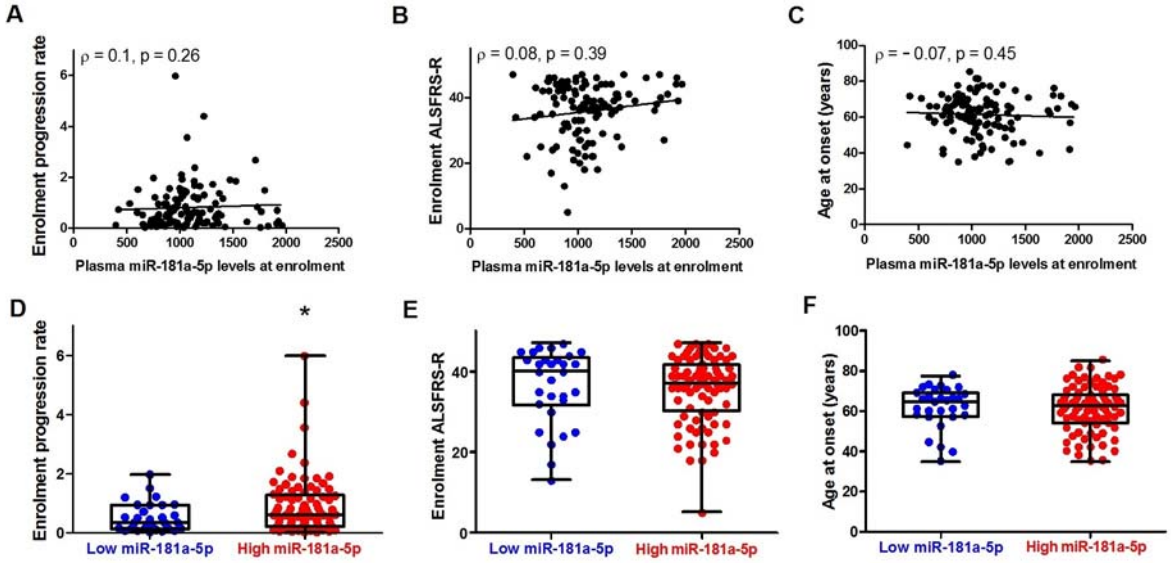
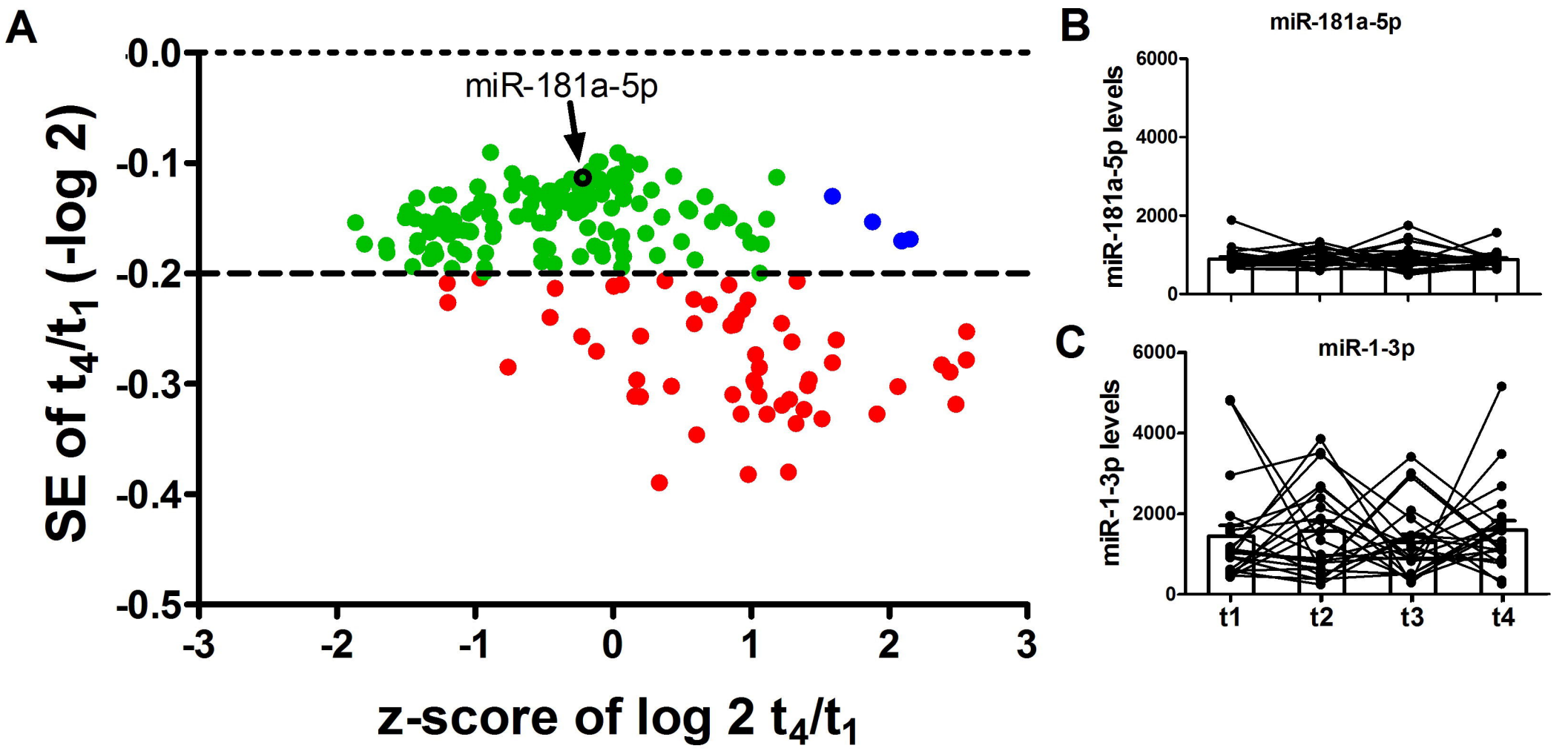
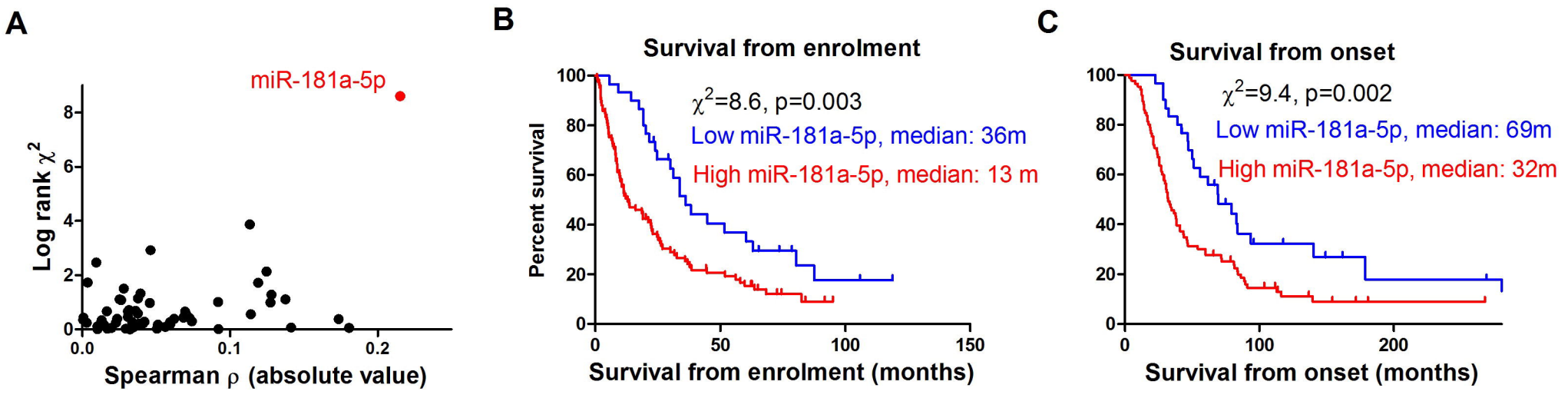
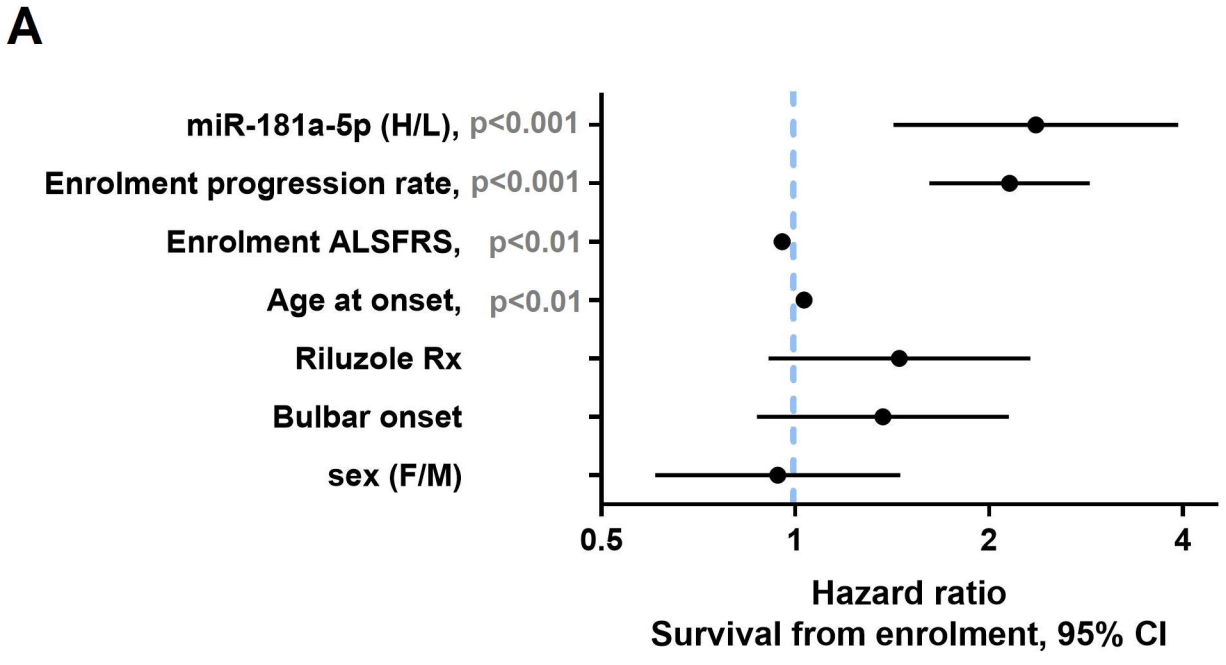
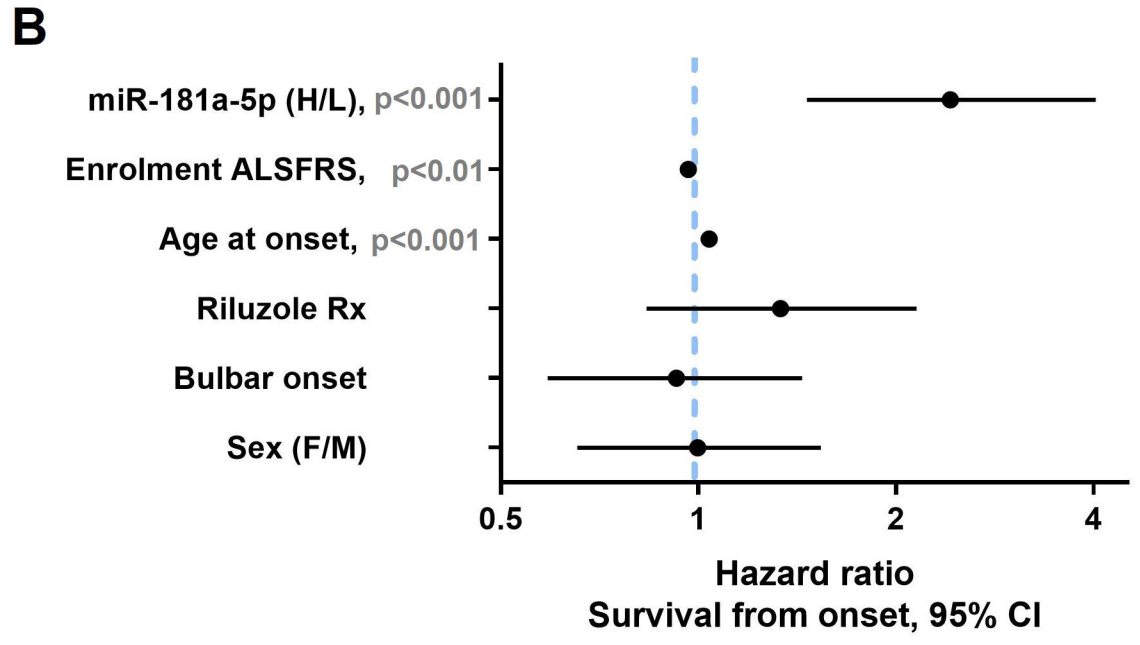


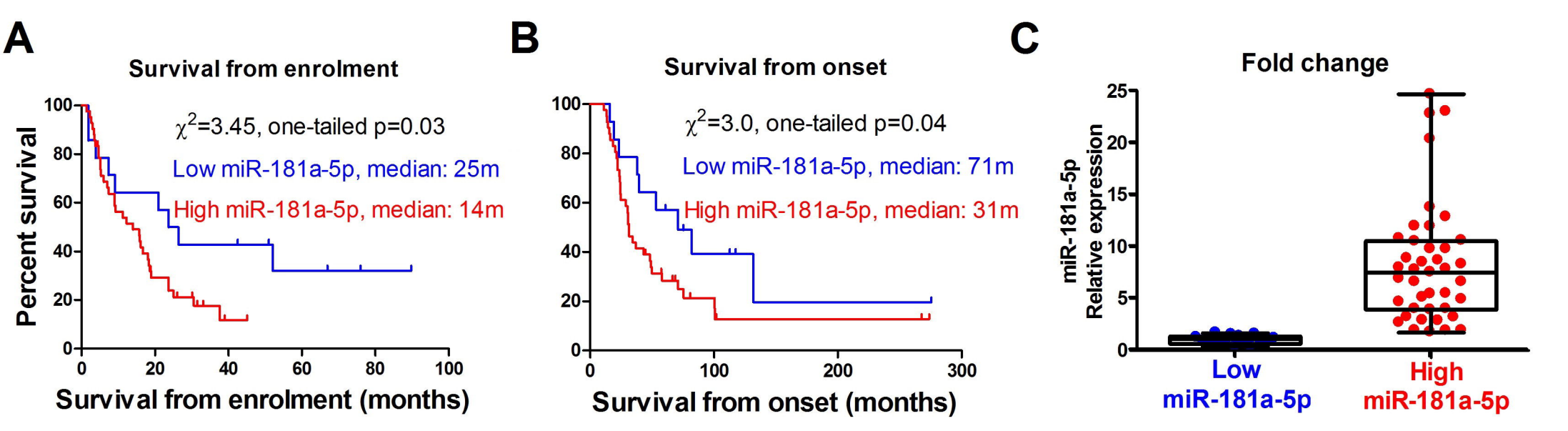
Figure S7. Correlation of miR-181a-5p levels at enrolment with phenotypic properties, including **(A))** progression rate at enrolment **(B)** enrolment ALSFRS-score, and **(C)** patient age at disease onset. Differences in these properties between low and high miR-181a bins **(D-F)**. *p<0.05 high vs low, Mann-Whitney U test.


A
miR-181a-5p
SE of t_4/t_1 (-log 2)
0.0
-0.1
-0.2
-0.3
-0.4
-0.5
-3
-2
-1
0
1
2
3
z-score of log 2 t_4/t_1
B
miR-181a-5p
miR-181a-5p levels
6000
4000
2000
0
C
miR-1-3p
miR-1-3p levels
6000
4000
2000
0
t1
t2
t3
t4


A
miR-181a-5p
Log rank χ^2
Spearman ρ (absolute value)
B
Survival from enrolment
χ^2=8.6, p=0.003
Low miR-181a-5p, median: 36m
High miR-181a-5p, median: 13 m
Percent survival
Survival from enrolment (months)
C
Survival from onset
χ^2=9.4, p=0.002
Low miR-181a-5p, median: 69m
High miR-181a-5p, median: 32m
Survival from onset (months)


A
miR-181a-5p (H/L), p<0.001
Enrolment progression rate, p<0.001
Enrolment ALSFRS, p<0.01
Age at onset, p<0.01
Riluzole Rx
Bulbar onset
sex (F/M)
0.5
1
2
4
Hazard ratio
Survival from enrolment, 95% CI


B
miR-181a-5p (H/L), p<0.001
Enrolment ALSFRS, p<0.01
Age at onset, p<0.001
Riluzole Rx
Bulbar onset
Sex (F/M)
0.5
1
2
4
Hazard ratio
Survival from onset, 95% CI


A
Survival from enrolment
χ^2=3.45, one-tailed p=0.03
Low miR-181a-5p, median: 25m
High miR-181a-5p, median: 14m
Percent survival
Survival from enrolment (months)
B
Survival from onset
χ^2=3.0, one-tailed p=0.04
Low miR-181a-5p, median: 71m
High miR-181a-5p, median: 31m
Survival from onset (months)
C
Fold change
miR-181a-5p Relative expression
Low miR-181a-5p
High miR-181a-5p