

Circulating microRNAs are related to cognitive domains in the general population

Authors

Konstantinos Melas^a, Valentina Talevi^a, Rika Etteldorf^a, Santiago Estrada^{a,b}, Dennis M. Krüger^c, Tonatiuh Pena^{c,d}, N. Ahmad Aziz^{a,e}, Andre Fischer^{c,f,g}, Monique M.B. Breteler^{a,h}

Affiliations

^a Population Health Sciences, German Centre for Neurodegenerative Diseases (DZNE), Venusberg-Campus 1/99, 53127 Bonn, Germany

^b AI in Medical Imaging, German Centre for Neurodegenerative Diseases (DZNE), Venusberg-Campus 1/99, 53127 Bonn, Germany

^c Department for Epigenetics and Systems Medicine in Neurodegenerative Diseases, German Center for Neurodegenerative Diseases, Von-Siebold-Str. 3a, 37075 Göttingen, Germany

^d Bioinformatics Unit, German Centre for Neurodegenerative Diseases, Von-Siebold-Str. 3a, 37075 Göttingen, Germany

^e Department of Neurology, Faculty of Medicine, University of Bonn, Venusberg-Campus 1/80, 53127 Bonn, Germany

^f Department for Psychiatry and Psychotherapy, University Medical Center Göttingen, Von-Siebold-Str. 5, 37075 Göttingen, Germany

^g Cluster of Excellence MBExC, University of Göttingen & University Medical Center Goettingen, Robert-Koch-Str. 40, 37075 Göttingen, Germany

^h Institute for Medical Biometry, Informatics and Epidemiology (IMBIE), Faculty of Medicine, University of Bonn, Venusberg-Campus 1/11, 53127 Bonn, Germany

Supplementary materials

Supplementary Methods S1: MicroRNA and gene expression sequencing and analysis

Supplementary Methods S2: Weighted gene co-expression network analysis

Supplementary Methods S3: MicroRNA target genes and functional enrichment analysis

Supplementary Table S1: Previous studies that identified circulating microRNAs related to cognition, AD or MCI

Supplementary Figure S1: Construction of weighted gene co-expression network

Supplementary Figure S2: Association of microRNA expression with cognitive scores in Model 2.

Supplementary Figure S3: Distribution of normalized counts for cognition-related microRNAs

Supplementary Figure S4: Neuropsychiatric and neurodegenerative disease sensitivity analysis.

Supplementary Figure S5: Stratified analysis for microRNAs inconsistently associated with cognition in our study and in two previous studies

Supplementary Figure S6: Predicted and downregulated microRNA target genes

References

Supplementary Methods 1

MicroRNA sequencing and analysis

Sequencing libraries for microRNA were prepared using NEBNext® small RNA library preparation kit according to manufacturer's instructions. Quality control of the sequencing was evaluated through FastQC v0.11.9 [1]. Trimmomatic v.0.39 software [2] was used to remove sequencing adapters, low-quality score reads, and reads shorter than 18 base pairs. The alignment and the quantification of microRNAs was performed using the mirdeep2 version 2.0.1.2 software package [3]. Briefly, following the trimming of low-quality score reads, the sequencing reads were aligned to the Human Genome GRCh38.p13 provided by Ensembl, and the expression of known microRNAs was determined by mapping the sequencing reads to predefined human precursor and mature microRNA sequences retrieved from miRBase v22.1 [4].

Gene expression sequencing and analysis

RNA integrity and quantity were assessed through the tapestation RNA assay on a TapeStation4200 instrument (Agilent). 750 ng of total RNA was used to generate NGS libraries for total RNA sequencing using the TruSeq stranded total RNA kit (Illumina) following manufacturer's instructions with Ribo-Zero Globin reduction. We checked library size distribution using D1000 on a TapeStation4200 instrument (Agilent) and quantified the libraries via Qubit HS dsDNA assays (Invitrogen). We clustered the libraries at 250 pM final clustering concentration on the NovaSeq6000 using S2 v1 chemistry (Illumina) in XP mode and sequenced paired-end 2*50 cycles before demultiplexing using bcl2fastq2 v2.20. Quality control of the sequencing was evaluated through FastQC v0.11.9. Following the trimming of low-quality score reads using Trimmomatic v.0.39 software, sequencing reads were aligned to the human reference genome GRCh38.p13 from Ensembl using STAR v2.7.1 [5]. The count matrix was generated with STAR –quantMode GeneCounts using the human gene annotation version GRCh38.101.

Supplementary methods 2

Construction of weighted gene co-expression network

To construct the WGCNA dendrogram, we first calculated pair-wise biweight midcorrelations of microRNA expression. We then chose a soft thresholding power of 7, which satisfied the scale-free topology criterion of $R^2 > 0.9$ (**Supplementary Figure S1**), and used a signed network to construct an adjacency matrix. Next, we converted the adjacency matrix to a topological matrix, using the topological overlap measure (TOM), and generated a hierarchical clustering tree with 1-TOM as the distance and using the UPGMA algorithm (method="average"). Using the Hybrid Dynamic Tree Cut method and setting a minimum module size of 5, we thus identified 16 microRNA modules. We set the *deepsplit* parameter to 4, to increase the overall number of clusters and avoid over-clustering.

Identification of module eigengenes and hub microRNAs

WGCNA permits users to reduce the expression of genes in each module to a single value. This value, called the module (eigen)expression, is defined as the first principal component

derived from Principal Component Analysis (PCA), and is representative of overall microRNA expression within the module. Using this method, we calculated the eigenexpression of each of the microRNA modules identified in our data. Another useful tool offered by WGCNA is the identification of *hub genes* within each module, which are seen as module representatives, or key drivers of module expression and its relation to traits. We defined *hub microRNAs* as microRNAs with *high module membership* and *high significance*, that is, microRNA expression had to be *a)* highly correlated with module eigenexpression (Pearson's $R \geq 0.8$ between microRNA expression and module eigenexpression) and *b)* significantly associated with the same trait as module expression.

Supplementary Methods 3

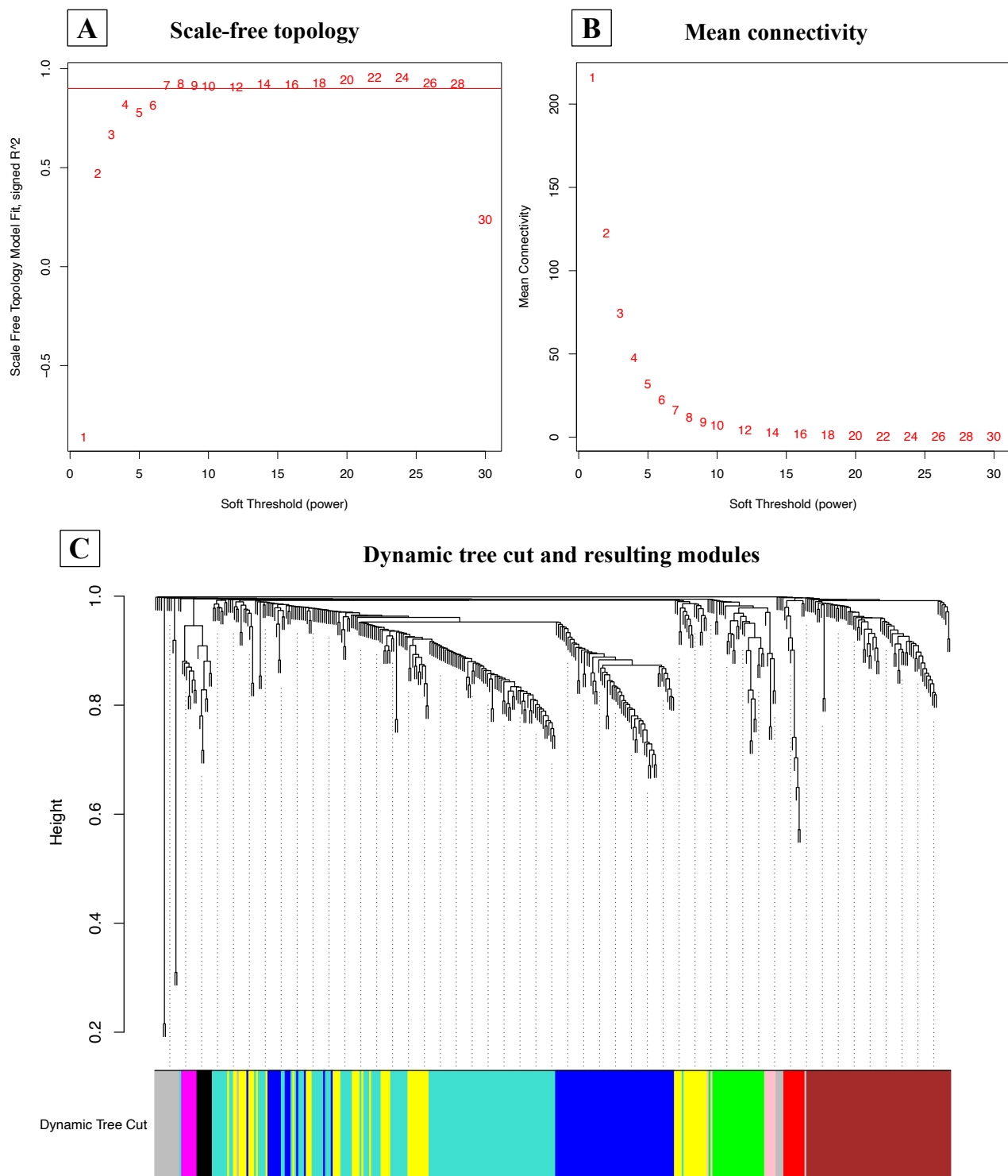
Identification of microRNA target genes

To identify genes that could be regulated by microRNAs, we examined the association between the expression of microRNAs and their potential target genes. We utilized the *multimir* R package (v.1.12.0) [6] to obtain a list of putative target genes for each microRNA from three online databases: experimentally validated microRNA-target gene interactions were obtained from the MirTarBase [7] database, while predicted microRNA-target gene interactions were obtained from the TargetScan [8] and miRDB [9,10] databases. To ensure the inclusion of all putative microRNA-target genes, the databases were interrogated for all predicted target genes, regardless of database prediction scores, and we employed the union of the target genes obtained from the 3 databases. Subsequently, we constructed separate regression models for each microRNA (independent variable) and their putative target genes (dependent variable), adjusting for age, sex, and blood cell counts. Lastly, we filtered for negatively associated target genes (linear regression beta coefficient < 0 and P value $\leq .05$), based on the assumption that microRNAs would downregulate the expression of their target genes.

Functional enrichment analysis

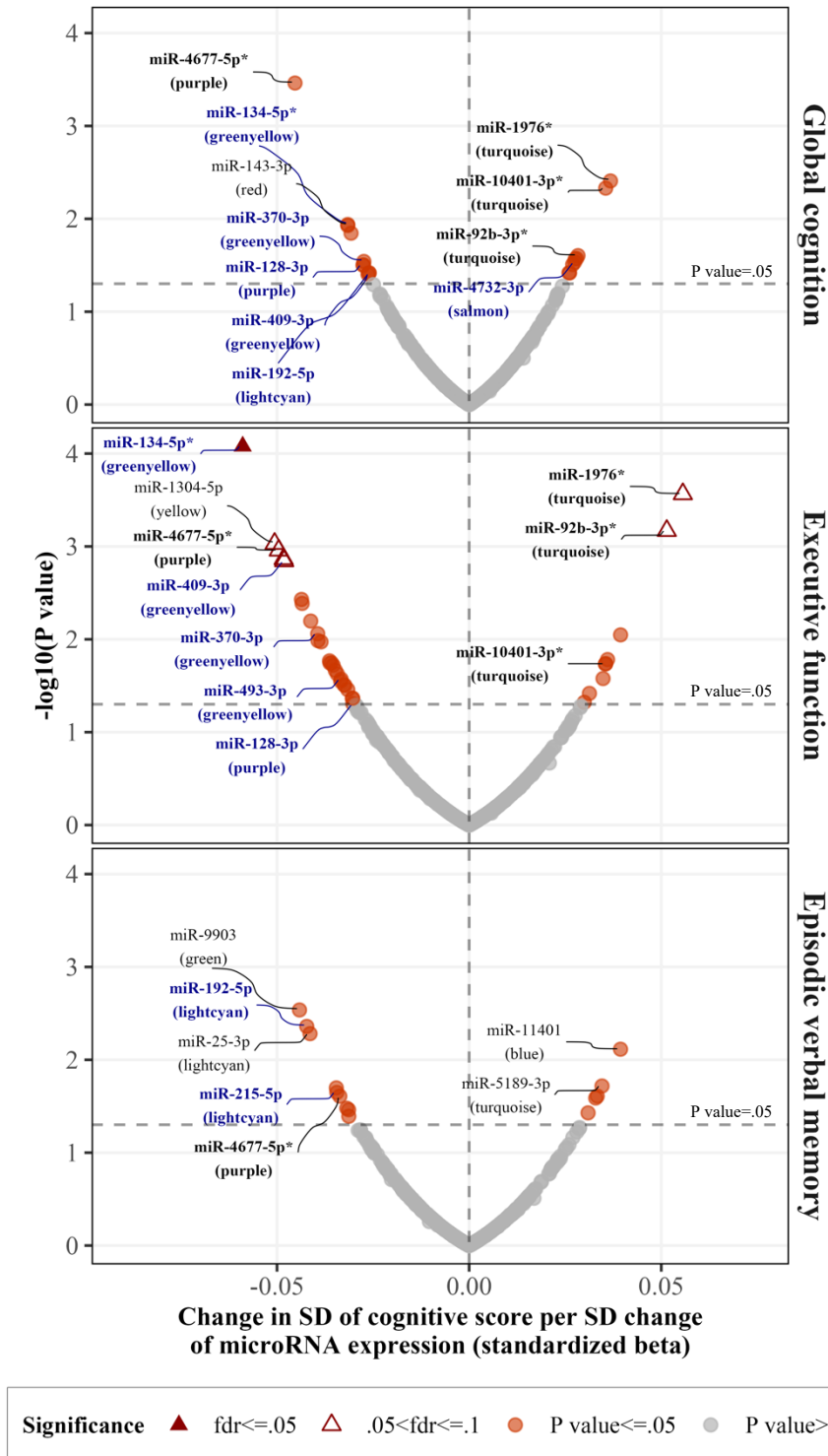
For the identified target genes we performed a pathway enrichment analysis using the *clusterProfiler* (v. 3.18.1) R Bioconductor package [11]. As we assume that microRNAs in WGCNA modules jointly regulate biological functions, we pooled the target genes of all the microRNAs belonging to the same module. We then conducted an overrepresentation analysis per module, using enrichment terms obtained from the Gene Ontology: Biological Processes (Gene Ontology) [12] database. The universe (background genes) of the overrepresentation analysis was set to the genes measured in our study ($n=11,325$). As enrichment terms from Gene Ontology tend to be redundant, similar terms were collapsed using the *rrvgo* (v. 1.2.0) R Bioconductor package [13]. Specifically, a semantic similarity matrix of Gene Ontology terms was constructed and similar terms were grouped, selecting the largest term (as defined by the number of genes in each enriched term) to represent each group. A small similarity threshold of 0.6 was selected to group only highly redundant terms, and the rest of the parameters were left at default settings. The P values of the terms within each group were pooled using the Fisher method and multiple testing adjustment according to the Benjamini-Hochberg false discovery rate method was performed for the pooled P values.

Supplementary Table S1: Previous studies that identified circulating microRNAs related to cognition, AD or MCI				
Study title	Study type	Sample size	MicroRNAs identified*	Reference
Genome-wide profiling of circulatory microRNAs associated with cognition and dementia.	Population-based, cross-sectional	n=1615	<u>miR-150-5p</u> , <u>miR-19b-3p</u> , <u>miR-20b-5p</u> , <u>miR-122-5p</u> , <u>miR-92a-3p</u> , <u>miR-4732-3p</u> , <u>miR-146a-5p</u> , <u>miR-30a-5p</u> , <u>miR-23b-3p</u> , <u>let-7c-5p</u> , <u>miR-4449</u> , <u>miR-324-3p</u> , <u>miR-6780b-5p</u> , <u>miR-2861</u> , <u>miR-4800-5p</u> , <u>miR-4688</u> , <u>miR-6800-3p</u> , <u>miR-185-5p</u> , <u>miR-3155b</u> , <u>miR-372-3p</u> , <u>miR-6126</u> , <u>miR-8071</u> , <u>miR-4697-5p</u> , <u>miR-7106-5p</u> , <u>miR-187-3p</u> , <u>miR-1322</u> , <u>miR-1275</u> , <u>miR-3141</u> , <u>miR-8078</u> , <u>miR-197-5p</u> , <u>miR-6870-5p</u> , <u>miR-6738-5p</u> , <u>miR-4539</u> , <u>miR-4522</u> , <u>miR-34a-5p</u> , <u>miR-4534</u> (Identified in multi-marker analysis)	Yaqub et al, 2022 [15]
Extracellular microRNA and cognitive function in a prospective cohort of older men: The Veterans Affairs Normative Aging Study.	Population-based (male participants), cross-sectional and longitudinal	n=530	Cross-sectional analysis: <u>miR-148a-5p</u> , <u>miR-335-3p</u> Longitudinal analysis: <u>miR-28-3p</u> , <u>miR-152-3p</u> , <u>miR-30e-3p</u> , <u>miR-148b-3p</u> , <u>miR-326</u> , <u>miR-22-3p</u> , <u>miR-221-3p</u> , <u>miR-4732-3p</u> , <u>miR-652-3p</u> , <u>let-7b-5p</u> , <u>miR-342-5p</u> , <u>miR-574-5p</u> , <u>miR-21-5p</u> , <u>miR-320c</u> , <u>miR-340-5p</u> , <u>miR-194-5p</u> , <u>miR-6842-3p</u> , <u>miR-363-3p</u> , <u>miR-584-5p</u> , <u>miR-99a-5p</u> , <u>miR-191-5p</u> , <u>miR-329-3p</u> , <u>miR-215-5p</u> , <u>miR-181a-5p</u> , <u>let-7f-5p</u> , <u>miR-320b</u> , <u>miR-199a-5p</u> , <u>miR-125b-5p</u> , <u>miR-511-5p</u> , <u>miR-6852-5p</u> , <u>miR-431-5p</u> , <u>miR-493-5p</u> , <u>miR-337-3p</u>	Comfort et al, 2022 [14]
A microRNA signature that correlates with cognition and is a target against cognitive decline.	Population-based, mouse models, case-control	Population-based: n=132; Case-control: n=71 MCI patients, n=65 HC	<u>miR-181a-5p</u> , <u>miR-146a-5p</u> , <u>miR-148a-3p</u>	Islam et al, 2021 [16]
The interplay between micro RNAs and genetic liability to Alzheimer's Disease on memory trajectories in the general population.	Population-based, longitudinal	n = 548	<u>miR-19a-3p</u> , <u>miR-485-5p</u>	Van der Auwera et al, 2023 [17]
Association of Blood MicroRNA Expression and Polymorphisms with Cognitive and Biomarker Changes in Older Adults.	Population-based, cross-sectional	n=830	<u>let-7a-5p</u> , <u>let-7c-5p</u> , <u>let-7d-5p</u> , <u>miR-144-5p</u> , <u>miR-93-5p</u> , <u>miR-98-5p</u> , <u>miR-363-3p</u> (Identified in linear regression analysis)	Sadlon et al [18]
Differential expression of MicroRNAs in Alzheimer's disease: a systematic review and meta-analysis	Meta-analysis	AD: n=2,787 patients, n=2,726 HC; MCI: n=302 patients, n=408 HC	AD: <u>miR-103a-3p</u> , <u>miR-106a-5p</u> , <u>miR-107</u> , <u>miR-1306-5p</u> , <u>miR-146a-5p</u> , <u>miR-15a-5p</u> , <u>miR-15b-3p</u> , <u>miR-20a-5p</u> , <u>miR-27b-3p</u> , <u>miR-29b-3p</u> , <u>miR-30e-5p</u> , <u>miR-335-5p</u> , <u>miR-361-5p</u> , <u>miR-18b-5p</u> , <u>has-miR-221-3p</u> , <u>miR-31-5p</u> , <u>miR-424-5p</u> , <u>miR-582-5p</u> , <u>miR-9-5p</u> MCI: <u>miR-128-3p</u> , <u>miR-132-3p</u> , <u>miR-323a-3p</u> , <u>miR-134-3p</u> , <u>miR-874-3p</u> (microRNAs differentially expressed in blood with I ² <0.4)	Yoon et al, 2022 [19]
* Underlined microRNAs were measured in the Rhineland Study. Abbreviations: MCI, Mild Cognitive Impairment; HC, Healthy Control; AD, Alzheimer's Disease				

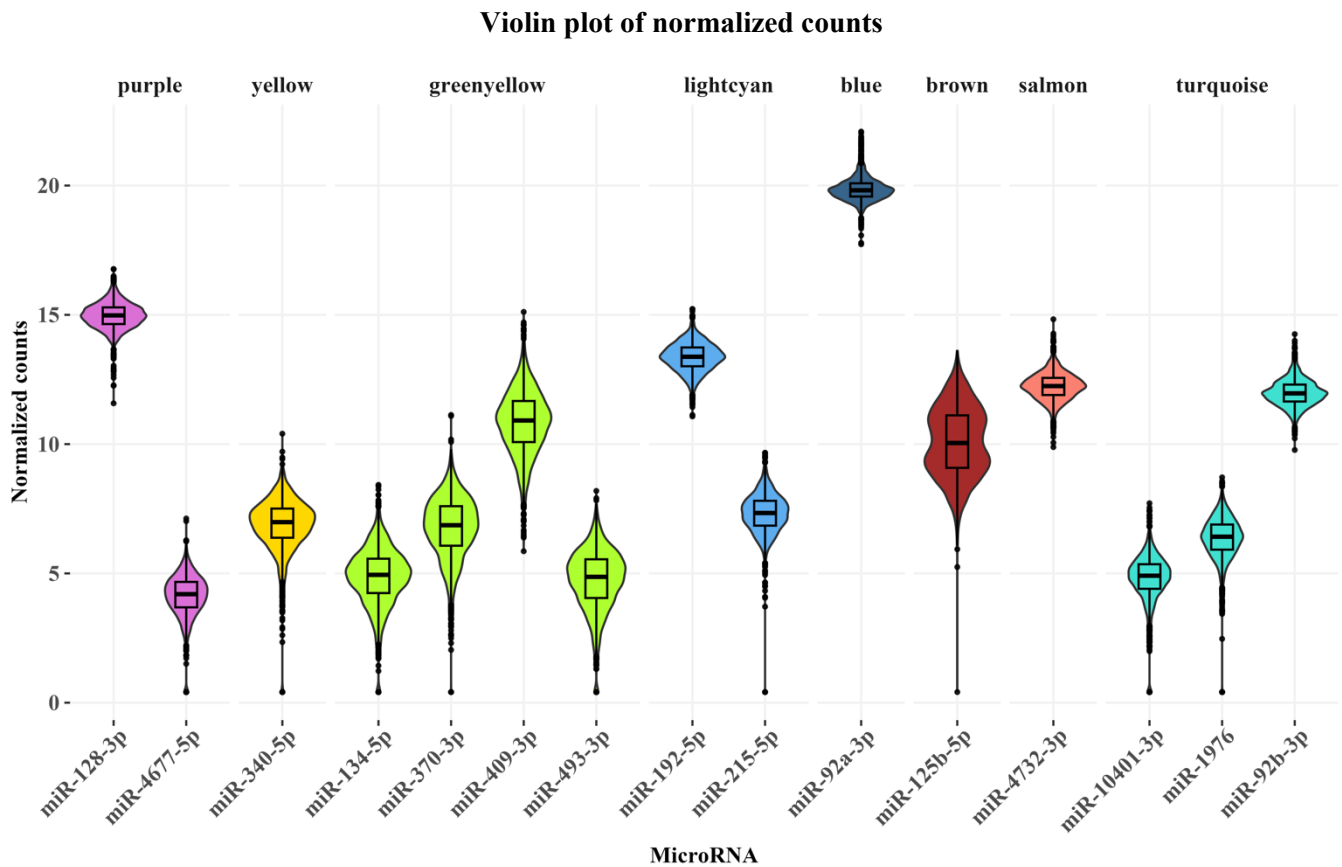


Supplementary Figure S1: Construction of weighted gene co-expression network
(A) Scale-free topology and **(B)** mean connectivity as a function of soft thresholding power.
(C) Implementation of the dynamic tree cut algorithm and distribution of microRNAs in color-coded clusters (modules). Resulting modules are shown as colors at the lower part of the plot.

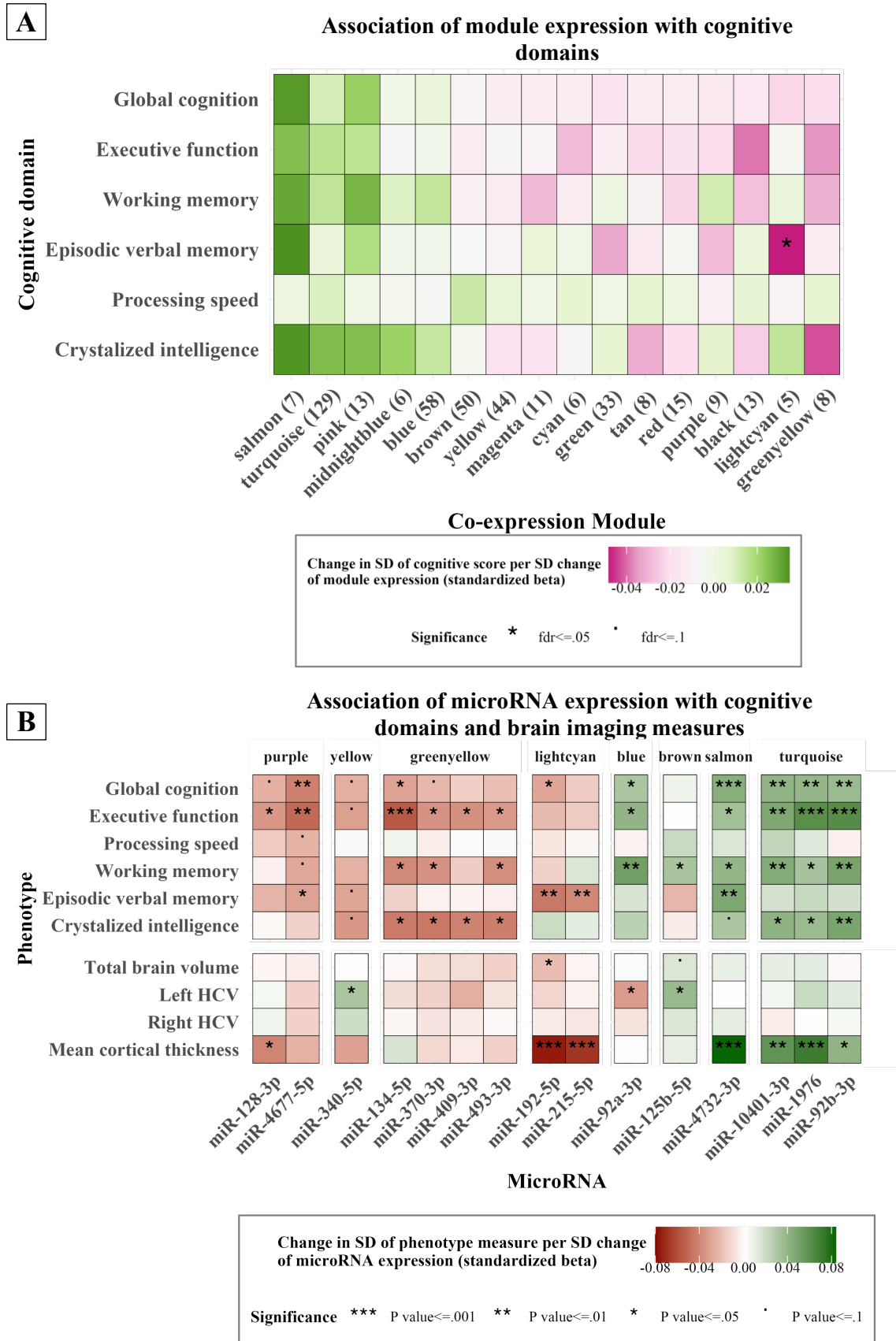
Association of microRNA expression with cognitive domains – Model 2



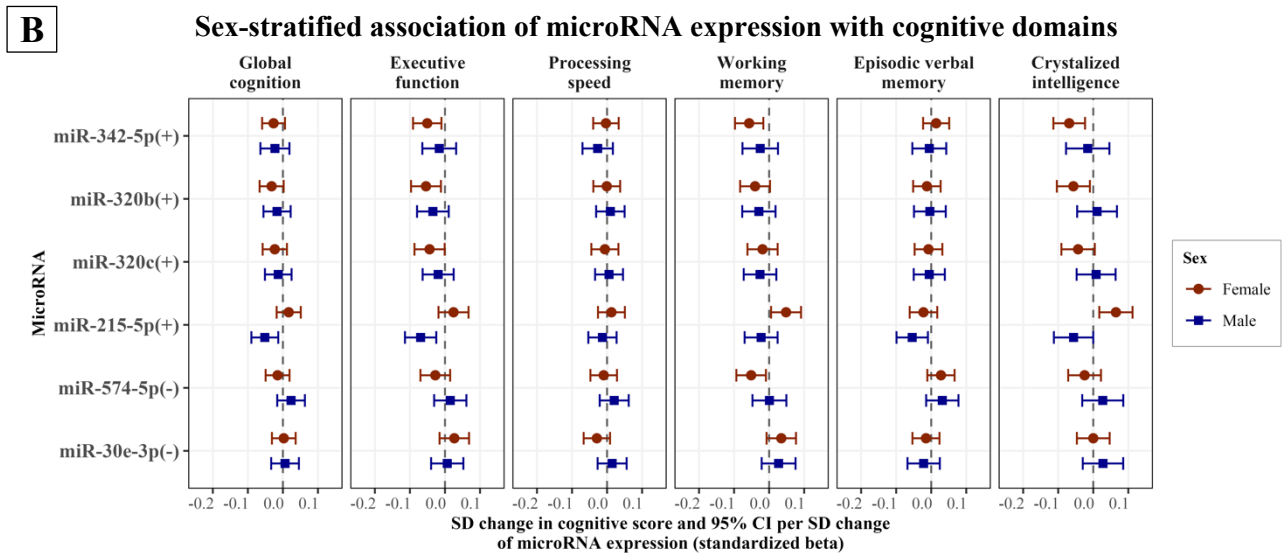
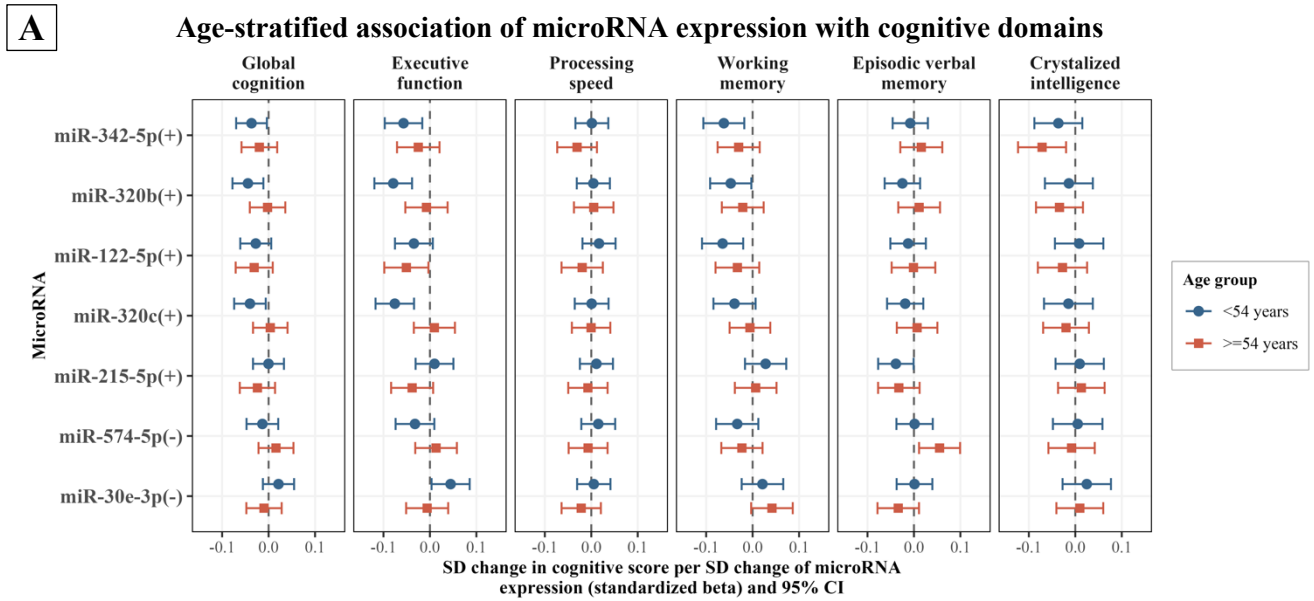
Supplementary Figure S2: Association of microRNA expression with cognitive scores in Model 2. Volcano plots of the association of all microRNAs profiled in our study with global cognition, executive function and episodic verbal memory scores. The P value shown in the y-axis has not been adjusted for multiple testing. Top microRNAs for each score are annotated and co-expression modules are shown in parentheses. MicroRNAs that reached statistical significance for at least one phenotype, after multiple testing correction ($fdr \leq 0.05$), are shown with triangles and have been marked with an asterisk (*). Hub microRNAs are annotated with blue color. Model 2 was used, which included as covariates age, sex, first language and educational level. Abbreviations: SD, Standard Deviation, fdr , false discovery rate



Supplementary Figure S3: Distribution of normalized counts for cognition-related microRNAs. The boxplots and violin plots show the distribution of DESeq2-normalized counts for the 15 identified cognition-related microRNAs, before adjusting for sequencing batch. MicroRNAs have been grouped according to the Weighted Gene Co-expression module they belong to.



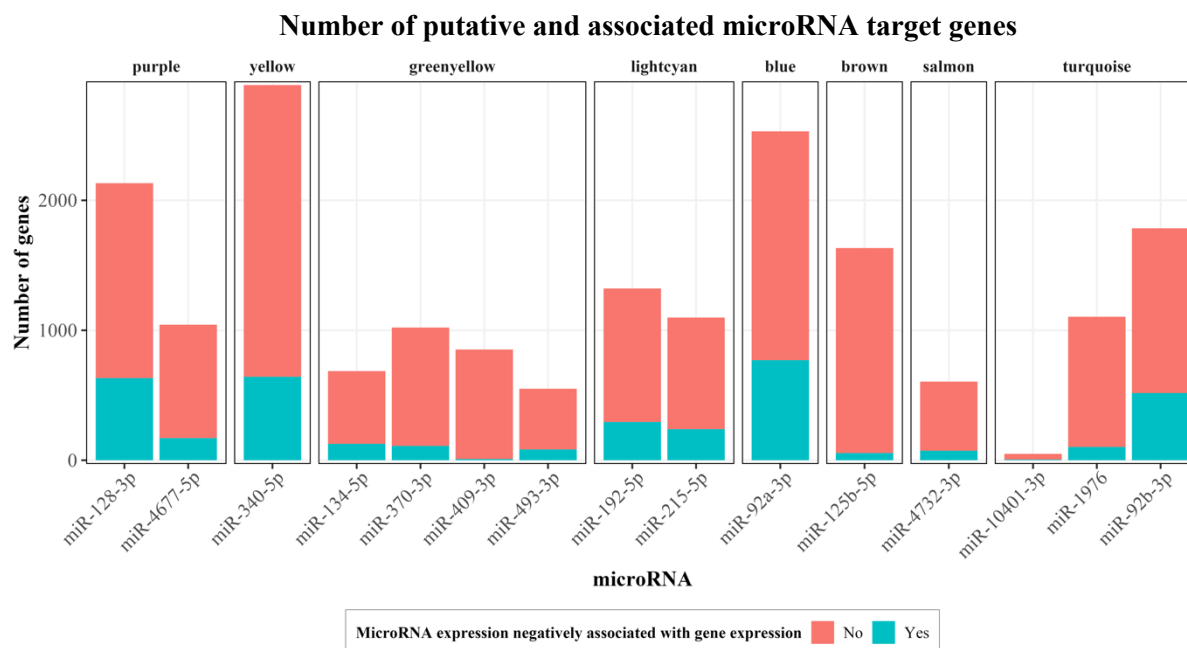
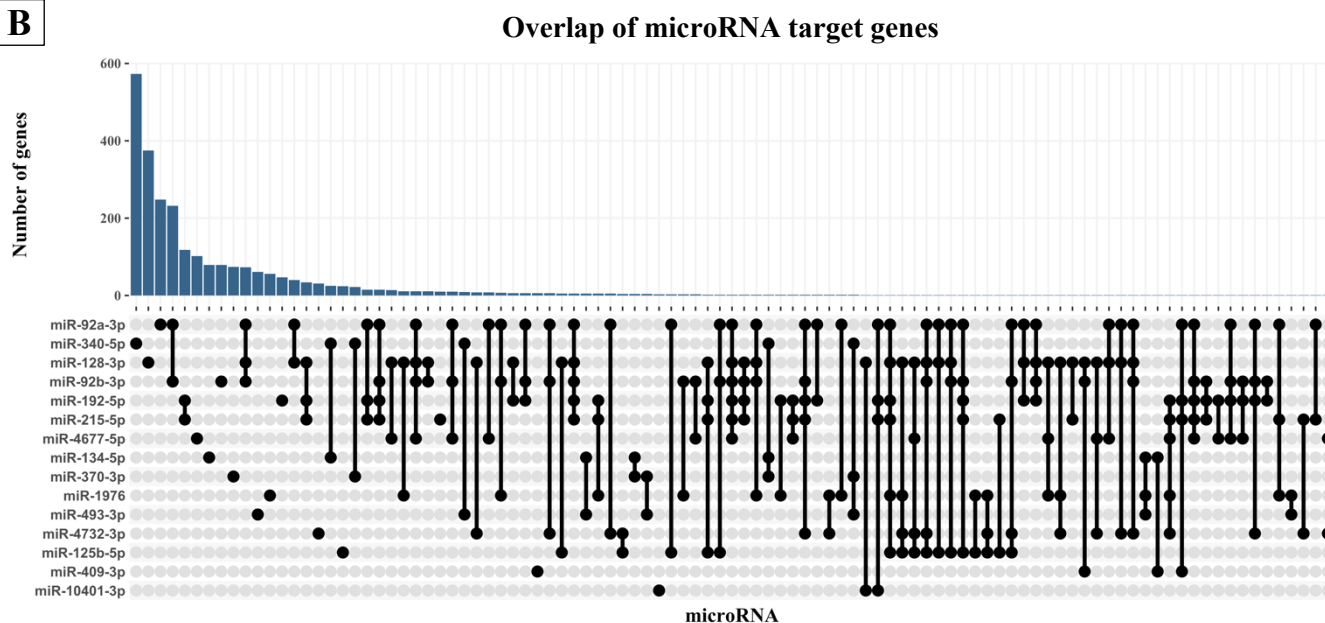
Supplementary Figure S4: Neuropsychiatric and neurodegenerative disease sensitivity analysis. Linear regression coefficients and significance levels for the association of (A) co-expression modules and (B) individual microRNAs with cognitive scores, after exclusion of participants who reported a physician diagnosis of dementia, Parkinson's disease, multiple sclerosis, stroke, and schizophrenia.



Supplementary Figure S5: Stratified analysis for microRNAs inconsistently associated with cognition in our study and in previous studies.

Seven microRNAs were associated with cognition in a different direction in our study and in two previous studies [14,15]. The forest plot shows the standardized regression betas and 95% confidence intervals the association of these microRNAs with cognitive scores after stratifying by (A) age and (B) sex. MicroRNAs are marked with “+” when their higher expression was associated with better cognitive function in the previous studies, and with “-” in the opposite case. Linear regression analyses were based on Model 1, which was adjusted for age.

Abbreviations: SD, Standard Deviation; CI, Confidence Interval

A**B**

Supplementary Figure S6: Predicted and downregulated microRNA target genes.

(A) The bar plot shows the number of genes that were predicted to be targeted by microRNAs, and the proportion of these predicted targets the expression of which was negatively associated with the expression of the targeting microRNA. MicroRNAs have been grouped by Weighted Gene Co-expression module. (B) The upset plot shows the overlap of negatively associated target genes per microRNA. The microRNAs are represented by dots. The lines linking the dots denote that the corresponding microRNAs have common negatively associated target genes. The number of these genes is shown in the bar plot above.

References

- [1] Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data 2010. <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/> (accessed November 29, 2022).
- [2] Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
- [3] Friedländer MR, Mackowiak SD, Li N, Chen W, Rajewsky N. miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. *Nucleic Acids Res* 2012;40:37–52. <https://doi.org/10.1093/nar/gkr688>.
- [4] Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res* 2019;47:D155–62. <https://doi.org/10.1093/nar/gky1141>.
- [5] Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 2013;29:15–21. <https://doi.org/10.1093/bioinformatics/bts635>.
- [6] Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, et al. The multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. *Nucleic Acids Res* 2014;42:e133. <https://doi.org/10.1093/nar/gku631>.
- [7] Huang H-Y, Lin Y-C-D, Li J, Huang K-Y, Shrestha S, Hong H-C, et al. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res* 2020;48:D148–54. <https://doi.org/10.1093/nar/gkz896>.
- [8] McGeary SE, Lin KS, Shi CY, Pham TM, Bisaria N, Kelley GM, et al. The biochemical basis of microRNA targeting efficacy. *Science* 2019;366. <https://doi.org/10.1126/science.aav1741>.
- [9] Chen Y, Wang X. miRDB: an online database for prediction of functional microRNA targets. *Nucleic Acids Res* 2020;48:D127–31. <https://doi.org/10.1093/nar/gkz757>.
- [10] Liu W, Wang X. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biol* 2019;20:18. <https://doi.org/10.1186/s13059-019-1629-z>.
- [11] Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation (Camb)* 2021;2:100141. <https://doi.org/10.1016/j.xinn.2021.100141>.
- [12] The Gene Ontology Consortium. The Gene Ontology resource: enriching a Gold mine. *Nucleic Acids Res* 2021;49:D325–34. <https://doi.org/10.1093/nar/gkaa1113>.
- [13] Sayols S. rrvgo: a Bioconductor package to reduce and visualize Gene Ontology terms. *Bioconductor*; 2020.
- [14] Comfort N, Wu H, De Hoff P, Vuppala A, Vokonas PS, Spiro A, et al. Extracellular microRNA and cognitive function in a prospective cohort of older men: The Veterans Affairs Normative Aging Study. *Aging (Albany NY)* 2022;14:6859–86. <https://doi.org/10.18632/aging.204268>.
- [15] Yaqub A, Mens MMJ, Klap JM, Weverling GJ, Klatser P, Brakenhoff JPJ, et al. Genome-wide profiling of circulatory microRNAs associated with cognition and dementia. *Alzheimers Dement* 2022. <https://doi.org/10.1002/alz.12752>.

- [16] Islam MR, Kaurani L, Berulava T, Heilbronner U, Budde M, Centeno TP, et al. A microRNA signature that correlates with cognition and is a target against cognitive decline. *EMBO Mol Med* 2021;13:e13659. <https://doi.org/10.15252/emmm.202013659>.
- [17] Van der Auwera S, Garvert L, Ameling S, Völzke H, Nauck M, Völker U, et al. The interplay between micro RNAs and genetic liability to Alzheimer's Disease on memory trajectories in the general population. *Psychiatry Res* 2023;323:115141. <https://doi.org/10.1016/j.psychres.2023.115141>.
- [18] Sadlon A, Takousis P, Evangelou E, Prokopenko I, Alexopoulos P, Udeh-Momoh CM, et al. Association of Blood MicroRNA Expression and Polymorphisms with Cognitive and Biomarker Changes in Older Adults. *J Prev Alzheimers Dis* 2024;11:230–40. <https://doi.org/10.14283/jpad.2023.99>.
- [19] Yoon S, Kim SE, Ko Y, Jeong GH, Lee KH, Lee J, et al. Differential expression of MicroRNAs in Alzheimer's disease: a systematic review and meta-analysis. *Mol Psychiatry* 2022;27:2405–13. <https://doi.org/10.1038/s41380-022-01476-z>.