Supplemental Information

Genome-wide, integrative analysis of circular RNA dysregulation and the corresponding circular RNA-microRNA-mRNA regulatory axes in autism

Yen-Ju Chen#, Chia-Ying Chen#, Te-Lun Mai#, Chih-Fan Chuang, Yu-Chen Chen, Sachin Kumar Gupta, Laising Yen, Yi-Da Wang, and Trees-Juen Chuang*

Supplemental Items:

Supplemental Table S1. The clinical features and the number of the identified circRNAs for the examined samples. (Excel file)

Supplemental Table S2. The identified circRNAs in the examined samples. (Excel file)

Supplemental Table S3. The 1,060 circRNAs used for the DE analysis. (Excel file)

Supplemental Table S4. The identified ASD-associated circRNA-miRNA and miRNA-mRNA regulatory axes. (Excel file)

Supplemental Table S5. The identified ASD-associated circRNA-miRNA-mRNA axes and the corresponding target genes. (Excel file)

Supplemental Table S6. Microarray analysis of the 171 target genes in the circARID1A-miR-204-3p-involved interactions in response to circARID1A knockdown, circARID1A overexpression, and miR-204-3p overexpression. (Excel file)

Supplemental Table S7. The primer, siRNA, miRNA mimic, miRNA inhibitor sequences and antibodies used in this study. (Excel file)

Supplemental Code. The scripts used in this study. (RAR file)

Supplemental Methods

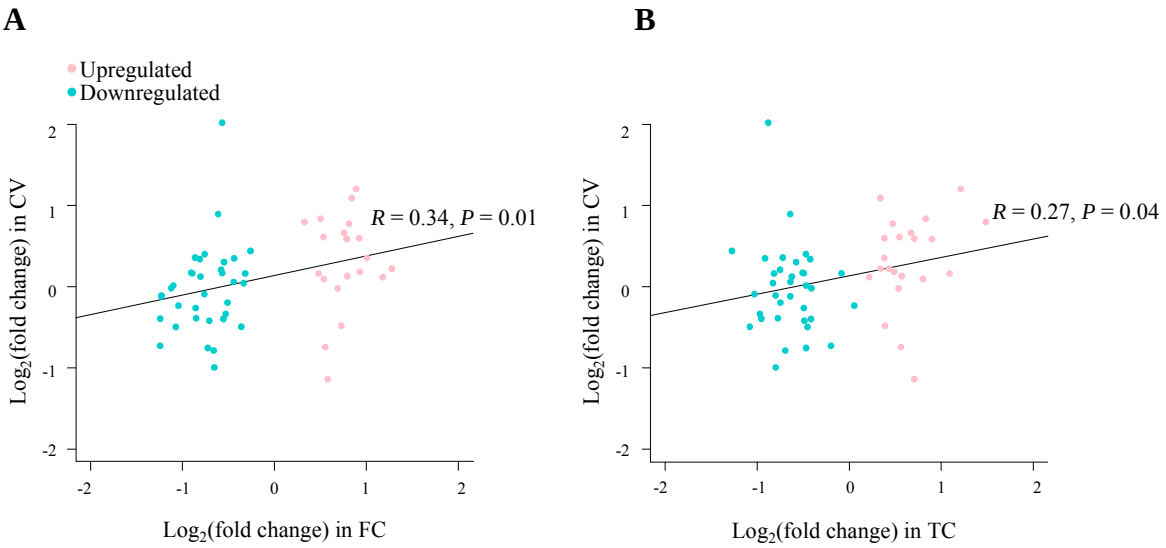
Weighted gene co-expression network analysis

First, pair-wise topology overlap (TO) between circRNAs was calculated with a soft threshold power of 9 to fit scale-free topology. Second, the network was constructed to be robust by performing 1000 rounds of bootstrapping strategy, and the TO matrix was computed for every resampling networks. The median of TOs was used to reconstruct a consensus TO matrix to avoid obtaining the module driven by outliers. Modules were identified using R package dynamicTreeCut with function cutreeDynamic, in which the paramenters are: method = "hybrid", deepSplit = 3, pamStage = T, pamRespectsDendro = T, minClusterSize = 10, and the expression of each module was summarized by the eigengene (ME), defined as the first principle component of all circRNAs in the module. To identify disease status, the Pearson correlations between MEs and several confounding factors, such as diagnosis, age, sex, brain region, RIN were calculated. Illustration of networks was plotted by the Cytoscape package (https://cytoscape.org/). Finally, to examine the similarity between FC and TC samples, and between ASD and CTL samples, the module preservation analysis was performed using WGCNA package with function modulePreservation. Each module was measured by the $Z_{summary}$ scores, where a module was regarded as "not preserved", "moderately preserved", and "highly preserved" if $Z_{summary}$ score < 2, 2 < $Z_{summary}$ score < 10, and $Z_{summary}$ score > 10, respectively. We found that the $Z_{summary}$ scores of all the three DE-modules were greater than or equal to two, indicating that the three modules were preserved across ASD and non-ASD control samples (Fig. 2D) and across FC and TC samples (Fig. 2E).

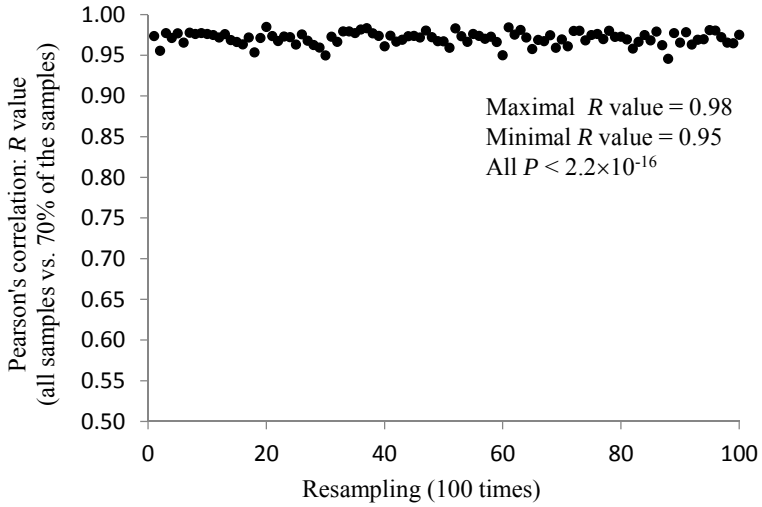
Identification of the targets of the DE-miRNAs

We identified the target genes of the 58 DE-miRNAs using the procedures illustrated in Supplemental Figure S4. Of the 58 DE-miRNAs, 37 are well-annotated and 21 are newly identified by Wu et al. (Wu et al. 2016) For the 37 well-annotated miRNAs, which can be found in the Ingenuity Pathway Analysis (IPA) package (Kramer et al. 2014) or databases with collection of experimentally supported miRNA-target interactions such as DIANA-TarBase (Karagkouni et al. 2018) and starBase (Li et al. 2014), we searched for targets of these miRNAs based on the MicroRNA Target Filter in the IPA package (QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis), DIANA-TarBase (version 8; obtained by request from the TarBase v8 administration team), and starBase (version 2.0; http://starbase.sysu.edu.cn/). We considered the miRNA-mRNA interactions if they satisfied one of the three criteria: (1) they were TargetScan-predicted (Agarwal et al. 2015) events with IPA filter score < -0.16 and also identified by Wu et al. (Wu et al. 2016); (2) they were experimentally supported events collected by Ingenuity® Expert Findings in IPA, which were manually curated by the IPA experts; and (3) they were experimentally supported events collected in DIANA-TarBase or starBase. Thus, we obtained 17,515 predicted and 26,217 experimentally-supported miRNA-mRNA axes, in which 9,004 axes were detected by both the prediction and experimental data. For the other 21 miRNAs, we identified targets of these miRNAs using the MirTarget algorithm with target prediction scores > 80 (the score was suggested by miRDB, in which the target prediction would be more likely to be real) via the miRDB (Wong and Wang 2015; Liu and Wang 2019) Web server at http://mirdb.org/custom.html, where the mature miRNA sequences of the 21 miRNAs were downloaded from the study of Wu et al. (Wu et al. 2016) For accuracy, we only considered the miRNA-mRNA interactions were also previously identified by Wu et al. (Wu et al. 2016), in which the interactions should satisfy one of the two criteria (the strongest target and the most conserved target criteria). We thus obtained 1,878 miRNA-mRNA axes based on the 21 new miRNAs.

Supplemental Figures

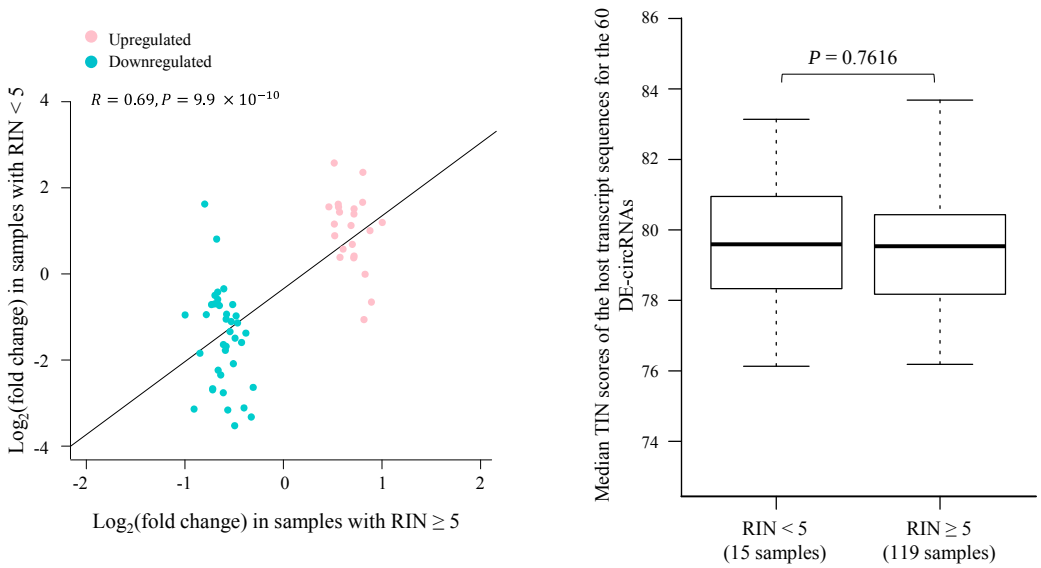


Supplemental Fig S1. Comparison of circRNA expression fold changes in the cortex (FC/TC) and cerebellum (CV): (A) FC vs. CV; (B) TC vs. CV. The black line represents the regression line between fold changes in the FC, TC, and CV for the 60 DE-circRNAs. The Pearson correlation coefficients (*R*) and *P* values are shown.

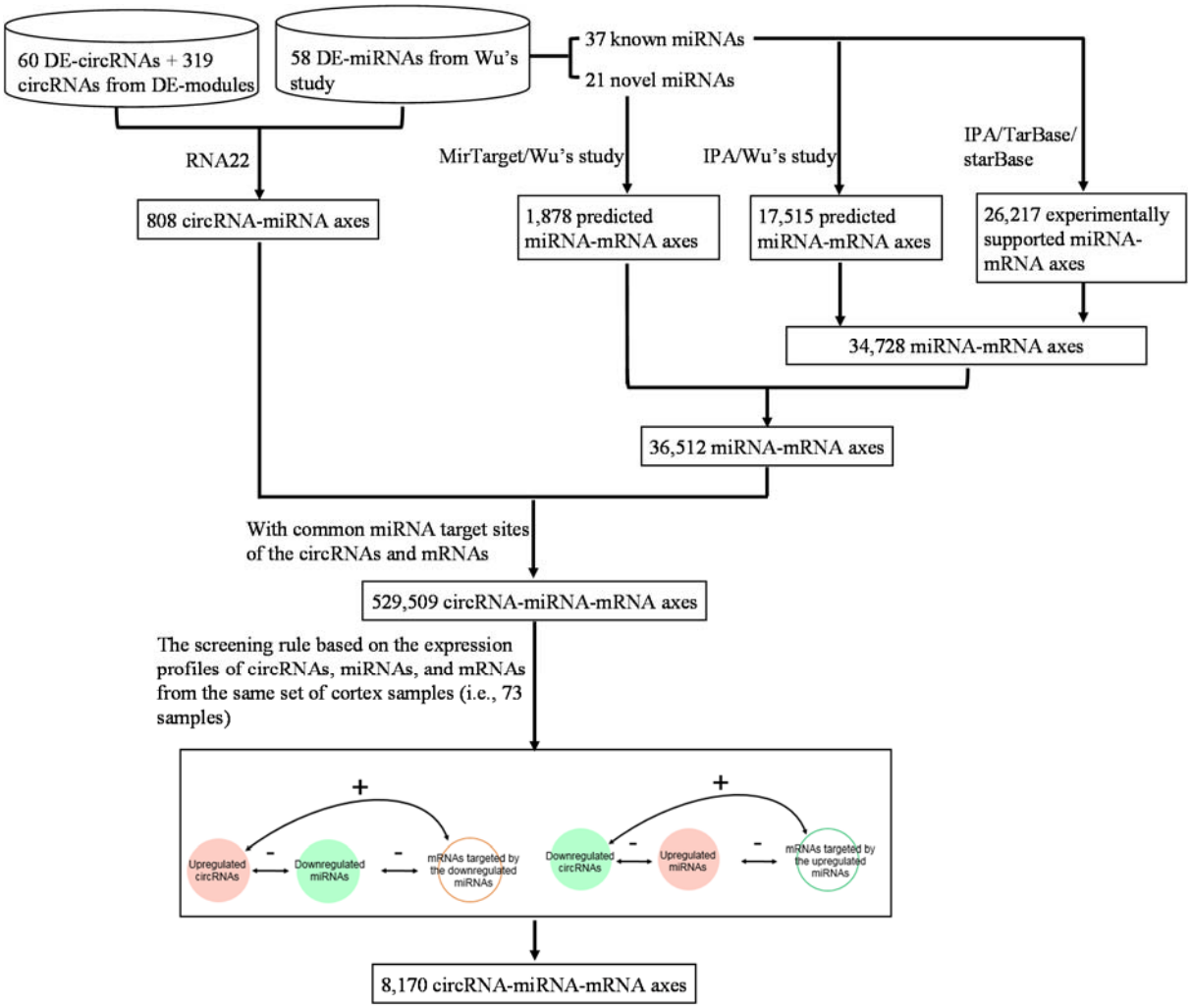


Supplemental Fig S2. Resampling analysis with 100 rounds of random sampling of 70% of the samples examined. The fold changes of the 60 DE-circRNAs for the resampled and the original sample sets were highly concordant between each other. The Pearson correlation coefficients (*R*) and *P* values are shown.

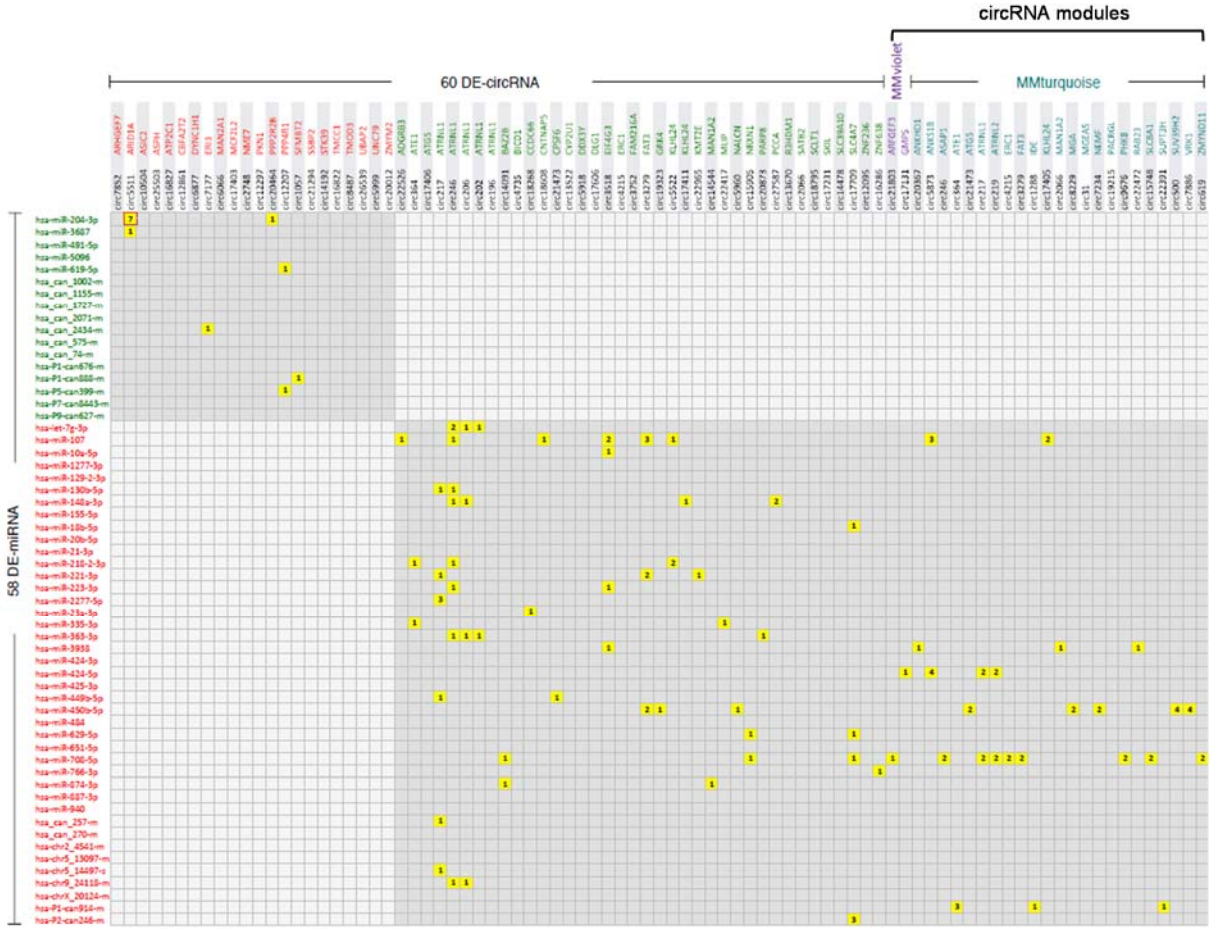
A **B**



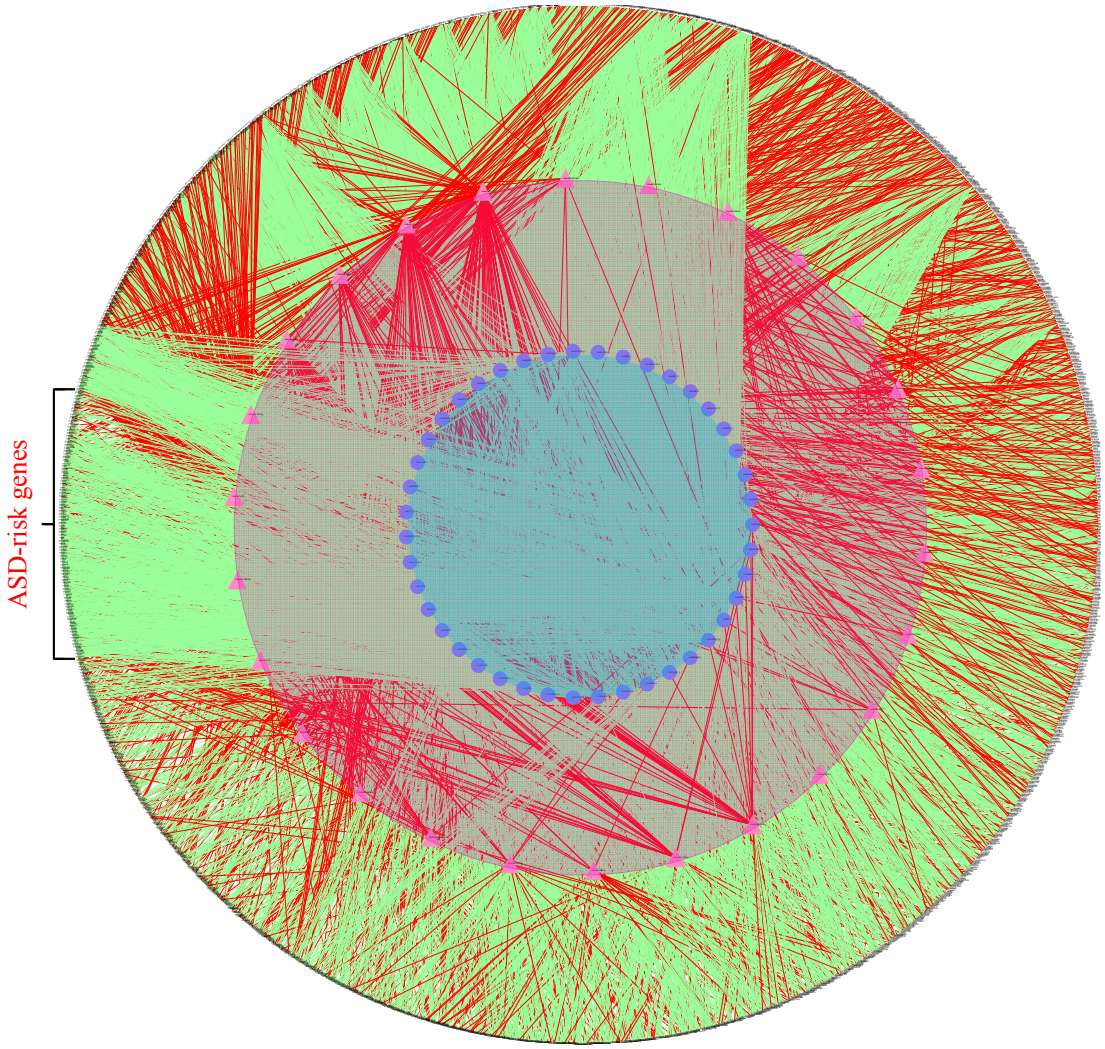
Supplemental Fig S3. Comparison of (A) circRNA expression fold changes and (B) the RNA quality of the host transcript sequences (measured by TIN) where the DE-circRNAs located for the cortex samples with RIN < 5 (15 samples) and RIN ≥ 5 (119 samples). For (A), the Pearson correlation coefficients (*R*) and *P* values are shown. For (B), *P* value was determined using two-tailed Wilcoxon rank-sum test.



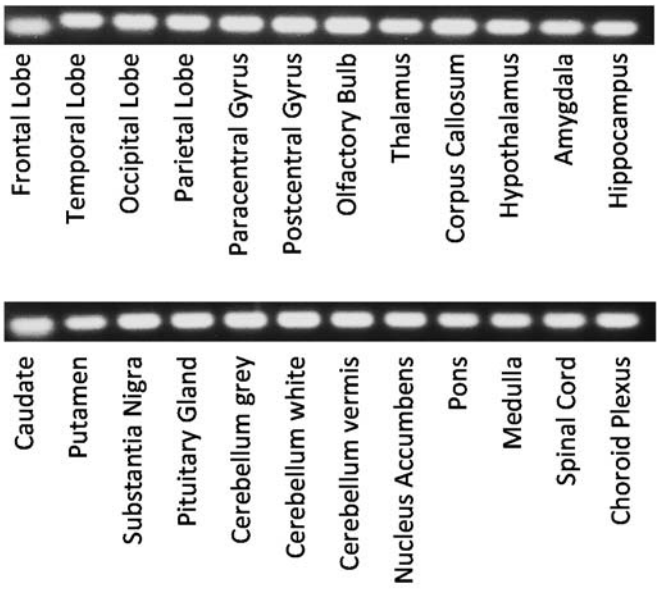
Supplemental Fig S4. Identification processes of the potential ASD-associated circRNA-miRNA-mRNA regulatory axes (see Methods).



Supplemental Fig S5. Numbers of the predicted binding sites of the DE-miRNAs in the circle sequences of the DE-circRNAs and circRNAs from the DE-modules. Upregulated and downregulated miRNAs/circRNAs are highlighted in red and green, respectively. Only the circRNA-miRNA regulatory axes with a significantly negative correlation (see also Fig. 3A) are shown. The red square highlights the DE-circRNA (circARID1A) that was predicted to have the greatest number of target sites of the DE-miRNAs (miR-204-3p).

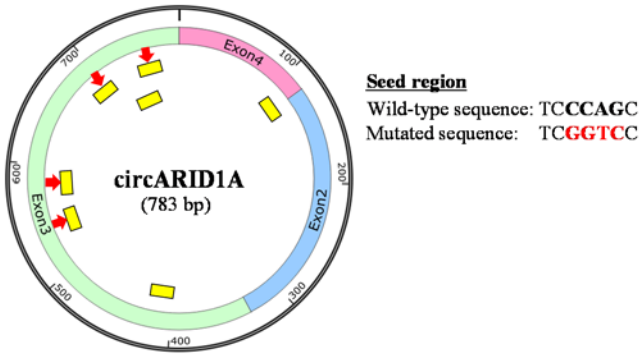


Supplemental Fig S6. The 7,814 downregulated circRNA-involved circRNA-miRNA-mRNA interactions plotted by the Cytoscape package.

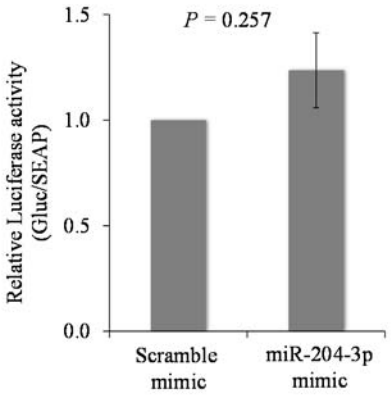


Supplemental Fig S7. RT-PCR analysis of circARID1A expression in 24 human brain regions.

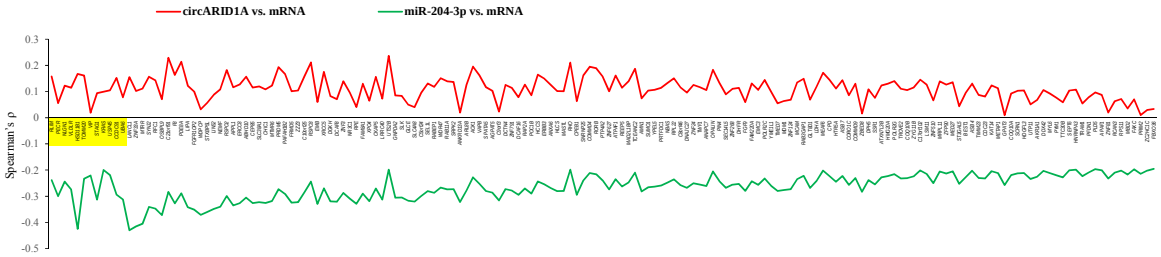
A



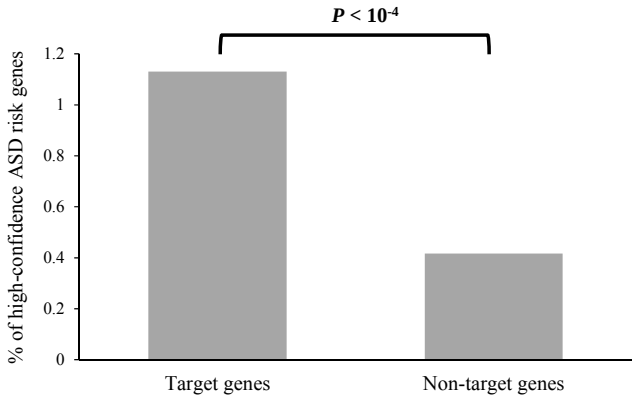
B



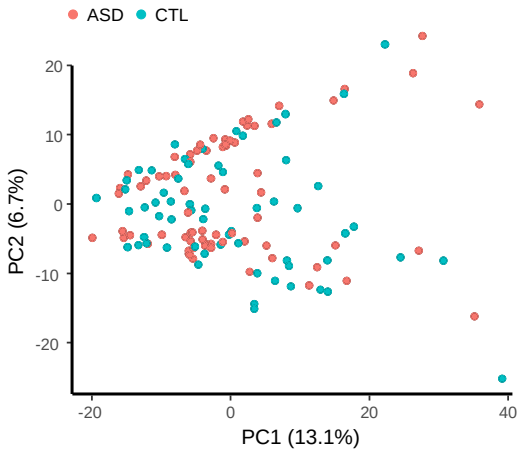
Supplemental Fig S8. Site-directed mutation of miR-204-3p binding sites in GLuc-circARID1A reporter construct. (A) The schematic diagrams displaying the four selected binding sites (the red arrows; see also Fig. 4A) in the circle sequence of circARID1A. The selected binding sites were then mutated from the luciferase-circARID1A reporter vector. The mutated sequence was also showed. (B) Luciferase report assay for the luciferase activity of GLuc-circARID1A-mutant in NHA cells co-transfected with miR-204-3p mimic and a scramble mimic. The qRT-PCR data are the means ±SD of three experiments. *P* value is determined using two-tailed *t*-test.



Supplemental Fig S9. The circRNA-mRNA and miRNA-mRNA correlations of expression profile between circARID1A and the 171 target mRNAs (red line) and between miR-204-3p and the 171 target mRNAs (green line), respectively. The circRNA-mRNA and miRNA-mRNA interactions of expression profile are determined using Spearman's correlation coefficients (ρ). Of the 171 targets, 12 are ASD risk genes (highlighted in yellow).



Supplemental Fig S10. Enrichment of high-confidence ASD risk genes for the targets of the identified circRNA-miRNA-mRNA interactions. The high-confidence ASD genetic risk factors were derived from a large dataset of whole-exome sequencing (35,584 ASD subjects; see Methods). The *P* value is determined using two-tailed Fisher's exact test.



Supplemental Fig S11. PCA based on the overall circRNA expression profiles of all identified circRNAs (1,060 circRNAs) from the cortex samples.

References

Agarwal V, Bell GW, Nam JW, Bartel DP. 2015. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**: e05005.

Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, Papadimitriou D, Kavakiotis I, Maniou S, Skoufos G et al. 2018. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res* **46**: D239-D245.

Kramer A, Green J, Pollard J, Jr., Tugendreich S. 2014. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**: 523-530.

Li JH, Liu S, Zhou H, Qu LH, Yang JH. 2014. starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res* **42**: D92-97.

Liu W, Wang X. 2019. Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. *Genome Biol* **20**: 18.

Wong N, Wang X. 2015. miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res* **43**: D146-152.

Wu YE, Parikshak NN, Belgard TG, Geschwind DH. 2016. Genome-wide, integrative analysis implicates microRNA dysregulation in autism spectrum disorder. *Nat Neurosci* **19**: 1463-1476.