Molecular Therapy
Nucleic Acids
Original Article

Whole-Transcriptome Analysis of APP/PS1 Mouse Brain and Identification of circRNA-miRNA-mRNA Networks to Investigate AD Pathogenesis

Nana Ma,[1] Jie Pan,[1] Xiaoyang Ye,[1] Bo Yu,[2,3] Wei Zhang,[1] and Jun Wan[1,4]

[1]Shenzhen Key Laboratory for Neuronal Structural Biology, Biomedical Research Institute, Shenzhen Peking University–The Hong Kong University of Science and Technology Medical Center, Shenzhen, Guangdong Province, China; [2]Shenzhen Key Laboratory for Translational Medicine of Dermatology, Biomedical Research Institute, Shenzhen Peking University–The Hong Kong University of Science and Technology Medical Center, Shenzhen, Guangdong Province, China; [3]Department of Dermatology, Peking University Shenzhen Hospital, Shenzhen, Guangdong Province, China; [4]Division of Life Science, The Hong Kong University of Science and Technology, Clear Water Bay Road, Kowloon, Hong Kong, China

Alzheimer's disease (AD) is one of the most common forms of dementia and is characterized by a progressive loss of cognition. A hallmark of AD is known to be the extensive distribution of neuronal tangles and amyloid plaques in the brain, but the molecular and cellular complexity of AD remains poorly elucidated, which limits the development of effective clinical treatments for AD. Accumulating evidence indicates that noncoding RNAs participate in AD-associated pathophysiology, but the details are largely unknown. Moreover, although recent studies have revealed a potential link between AD and circular RNA (circRNA)-associated competing endogenous RNA (ceRNA) networks, few genome-wide studies have identified putative circRNA-associated ceRNA pairs involved in AD. Here, we used deep RNA sequencing to systematically investigate circRNA-associated ceRNA mechanisms in the brain of AD model mice (APP/PS1). Our results identified 235, 30, and 1,202 significantly dysregulated circRNAs, microRNAs (miRNAs), and mRNAs, respectively, and we used the sequencing data to construct the most comprehensive circRNA-associated ceRNA networks to date in the APP/PS1 brain. Gene Ontology (GO) analysis revealed that the identified networks are involved in regulating AD development from distinct origins, such as from the dendrite (GO: 0030425) and the synapse (GO: 0045202). Following rigorous selection, the circRNA-associated ceRNA networks in this AD mouse model were discovered to be mainly involved in dendritic development and memory (*Sorbs2*) and mouse neural development (ALS2). This study presents the first systematic dissection of circRNA-associated ceRNA profiles in the APP/PS1 mouse brain, and the identified circRNA-associated ceRNA networks could provide insights that facilitate AD diagnosis and therapy in the future.

INTRODUCTION

Alzheimer's disease (AD) is a slowly progressing neurodegenerative disorder that begins with mild memory loss and culminates in severe impairment of executive and cognitive functions. Over the past decade, substantial progress has been made in identifying AD-related molecular and cellular processes, but the molecular mechanisms underlying AD pathogenesis remain largely unknown,[1] and the available pharmacological treatments for AD do not slow or stop the disease.[2] Therefore, further investigation of the mechanisms underlying AD is urgently required to enhance our understanding of the disease and to facilitate the development of effective therapeutic strategies.

Accumulating evidence indicates that noncoding RNAs, particularly microRNAs (miRNAs), long noncoding RNAs (lncRNAs), and circRNAs, participate in AD-associated pathophysiology, including the formation and development of β-amyloid (Aβ) plaques and neurofibrillary tangles, synaptic loss, and neuronal death.[3,4] circRNAs are evolutionarily conserved transcripts featuring covalently linked 5′ and 3′ ends that are derived from pre-mRNA back-splicing.[5] Because circRNAs lack 5′ and 3′ ends, their lifetimes typically range from hours to days or longer, making them considerably more stable than linear coding or noncoding mRNAs.[6] Thus, the biological functions of circRNAs might differ from those of other classes of RNAs, and the stability and specific expression of certain circRNAs have identified these RNAs as optimal candidates for biomarkers of aging and neurodegenerative diseases. circRNAs are enriched in neurons and synaptosomes,[7] and, intriguingly, fly and mouse studies have shown that brain circRNAs are regulated during

Received 21 July 2019; accepted 24 October 2019;
https://doi.org/10.1016/j.omtn.2019.10.030.

Correspondence: Jun Wan, Shenzhen Key Laboratory for Neuronal Structural Biology, Biomedical Research Institute, Shenzhen Peking University–The Hong Kong University of Science and Technology Medical Center, Shenzhen, Guangdong Province, China.
E-mail: wanj@ust.hk
Correspondence: Wei Zhang, Shenzhen Key Laboratory for Neuronal Structural Biology, Biomedical Research Institute, Shenzhen Peking University–The Hong Kong University of Science and Technology Medical Center, Shenzhen, Guangdong Province, China.
E-mail: zhangweispace@163.com

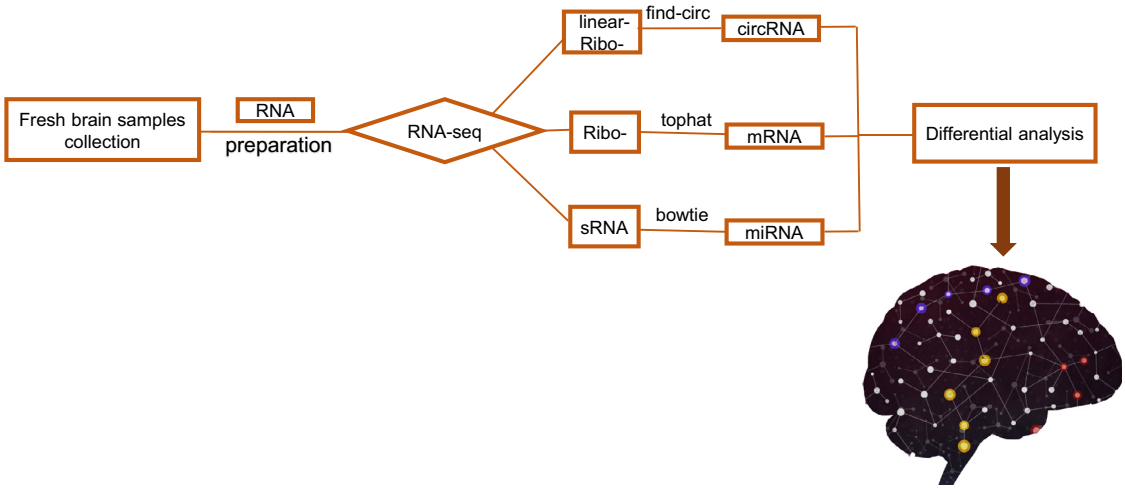


Figure 1. The Workflow of RNA-Seq
Details of the mRNA-seq, miRNA-seq, and circRNA-seq methods are described in Document S1.

aging, which strongly indicates potential roles for circRNAs in brain aging and aging-associated neurodegenerative diseases.[8,9]

A current model on circRNAs posits that these RNAs inhibit the functions of miRNAs by acting as miRNA sponges through the competing endogenous RNA (ceRNA) network.[10] These circRNA-associated ceRNA networks might contribute to several disease processes: Han et al.[11] determined that the circRNA circMTO1 acts as a sponge of miRNA-9 to suppress hepatocellular carcinoma progression. Zheng et al.[12] reported that TTBK2 circRNA promotes glioma malignancy by regulating miR-217. Huang et al.[13] found that the circRNA MYLK competitively binds to miRNA-29a-3p, and this results in increased expression of the target genes *DNMT3B*, *VEGFA*, and *ITGB1*, which are involved in the progression of bladder cancer. Fan et al.[14] summarized the functions of circRNAs and their role as a biomarker in cardiovascular diseases. Moreover, Peng et al.[15] reported that the circular RNA ZNF609 functions as a ceRNA to regulate AKT3 expression by sponging miR-150-5p in Hirschsprung's disease.

CDR1as (also known as CiRS-7: a circular RNA sponge bound to miR-7) and miR-7 are both reported to be associated with nervous system development and disease, and CDR1as dysfunction has been found to upregulate miR-7 expression and potentially lead to the downregulation of AD-related targets, including ubiquitin-protein ligase A.[16–18] The circRNA derived from *SRY* can act as a natural miRNA sponge and thus inhibit the activity of miR-138,[10] which can affect learning and memory by regulating acyl protein thioesterase 1.[19,20] These are the only reported examples thus far of circRNAs affecting the brain and the nervous system, but additional RNAs might participate in both normal brain function and nervous system disease.

The aforementioned findings raise the possibility that a favorable strategy for gaining insights into neurodegenerative disorders such as AD would be to comprehensively investigate the functions of circRNA-associated ceRNA networks. Elucidation of the potential link between AD and circRNA-associated ceRNA networks could suggest new strategies for combating this life-threatening disease. For the assessment of new potential treatments for AD, mouse models represent one of the most critical research tools. Thus, in this study, we used deep RNA sequencing (RNA-seq) to examine circRNA-associated ceRNA networks in the brain of APP/PS1 mice (which express APP695swe and PS1-dE9 mutations) and wild-type control mice at the 6- and 9-month-old stages. These time points were selected for the measurements because in this research model, Aβ can be detected in the mice when they are 6 months old, extracellular Aβ deposits in the cortex are apparent by 9 months of age, and synaptic transmission and long-term potentiation are clearly impaired in these mice when they are 9 months old.[21] RNA-seq is widely used to determine the differential gene expression profiles that underlie phenotypic differences,[22,23] and in this study, RNA-seq was used for the first time to identify circRNA-associated ceRNA networks in the APP/PS1 mouse model of AD (Figure 1); the data we obtained can serve as a useful resource in efforts to develop new therapeutic targets or novel diagnostics for AD.

RESULTS

Overview of circRNA-Seq Data

The sequencing generated 739,349,016 raw reads in total: 177,433,438 and 191,552,170 for wild-type and APP/PS1 6-month-old mice, and 185,516,382 and 184,847,026 for wild-type and APP/PS1 9-month-old mice, respectively. After poly(N)-containing, low-quality, and adaptor-containing reads were removed from the raw data, 724,361,150 clean reads remained: 173,777,328 and 186,369,844 for wild-type and APP/PS1 6-month-old mice, and 181,497,660 and 182,716,318 for wild-type and APP/PS1 9-month-old mice. The reference genome and gene model annotation files were downloaded from the genome website directly (ftp://ftp.ensembl.org/pub/release-93/

fasta/mus_musculus/dna/). An index of the reference genome was built using bowtie2 v.2.2.8, and paired-end clean reads were aligned to the reference genome by using Bowtie.[24] circRNAs were detected and identified using find_circ[6] and CIRI2,[25] and this yielded 15,713 circRNAs, which were used for subsequent analyses.

Overview of miRNA-Seq Data

Here, 99,137,837 raw reads were generated in total: 28,672,205 and 20,760,197 for wild-type and APP/PS1 6-month-old mice, and 23,016,993 and 26,688,442 for wild-type and APP/PS1 9-month-old mice. After removal of low-quality and adaptor sequences, 96,028,757 clean reads were obtained: 27,710,151 and 20,281,596 for wild-type and APP/PS1 6-month-old mice, and 22,393,001 and 25,644,009 for wild-type and APP/PS1 9-month-old mice. We filtered these results based on length (18–35 nt), and most selected reads were 22 nt in length in both groups. The selected reads were next mapped to the mouse reference sequence by using Bowtie,[24] and the mapping rates were 94.61% and 94.78% for APP/PS1 and wild-type mice, respectively. These mapped tags were annotated and classified based on alignment with noncoding small RNAs (rRNA, tRNA, small nuclear RNA [snRNA], and small nucleolar RNA [snoRNA]) in GenBank, repeat-associated RNA, exon- and intron-associated RNAs, and miRBase v.20.0, and we integrated miREvo[26] and miRDeep2[27] software to predict previously unidentified miRNAs. Ultimately, 1,411 mature miRNAs (1,312 known, 99 previously unknown) were detected, and these were used for subsequent analyses.

Overview of mRNA-Seq Data

A total of 1,145,853,722 raw reads were generated: 317,603,558 and 276,260,804 for wild-type and APP/PS1 6-month-old mice, and 286,203,120 and 265,786,240 for wild-type and APP/PS1 9-month-old mice. After discarding the reads including adapters, poly(N) > 10%, and any other potential contaminants, 1,119,187,680 clean reads were obtained: 311,741,312 and 266,033,410 for wild-type and APP/PS1 6-month-old mice, and 278,938,832 and 262,474,126 for wild-type and APP/PS1 9-month-old mice. Reference genome and gene model annotation files were downloaded directly from the genome website (ftp://ftp.ensembl.org/pub/release-93/fasta/mus_musculus/dna/). An index of the reference genome was built using bowtie2 v.2.2.8, and paired-end clean reads were aligned to the reference genome by using HISAT2 v.2.0.4.[28] HISAT2 was run with "–rna-strandness RF," other parameters were set as default, and the mapping rates were 91.53% and 91.74% for APP/PS1 and wild-type mice, respectively. We filtered out transcripts that were predicted to hold coding potential by the either/all function of CNCI (Coding-Non-Coding-Index) (v.2),[29] CPC (Coding Potential Calculator) (0.9-r2),[30] PfamScan (v.1.3),[31] and PhyloCSF (phylogenetic codon substitution frequency) (v.20121028),[32] and among these transcripts, 48,718 protein-coding transcripts were identified; these mRNAs were used for subsequent analyses.

Differential Expression Analysis: APP/PS1 versus Wild-Type

We identified 343 significantly dysregulated circRNA transcripts between the two groups at 6 months (Figure 2A; Table S1); 192 and 151 transcripts were upregulated and downregulated in APP/PS1 mice relative to their levels in wild-type mice, respectively. Moreover, at 9 months, 243 circRNA transcripts were significantly dysregulated: 141 and 102 transcripts were upregulated and downregulated in APP/PS1 mice relative to wild-type (Figure 2B; Table S2). Cluster analysis of the circRNAs expression was conducted, and a heatmap was constructed to visualize the results (Figure 2C). Next, based on transcripts per million (TPM) values, 36 significantly dysregulated miRNAs were identified between the two groups at 6 months: 12 miRNAs were upregulated and 24 were downregulated in APP/PS1 mice (Figure 2D; Table S3); at 9 months, 56 miRNAs were significantly dysregulated: 27 and 29 miRNAs were respectively upregulated and downregulated in APP/PSI mice (Figure 2E; Table S4). Cluster analysis of miRNA expression was performed, and a heatmap was generated (Figure 2F). Lastly, we used the value FPKMs (fragments per kilobase of exons per million fragments mapped) to estimate the expression levels of mRNA transcripts. At 6 months, 1,770 mRNA transcripts were significantly dysregulated, with 824 and 946 being respectively upregulated and downregulated in APP/PS1 mice (Figure 2G; Table S5); at 9 months, 1,678 mRNAs were significantly dysregulated, with 832 and 846 being upregulated and downregulated in APP/PS1 mice (Figure 2H; Table S6). Once again, cluster analysis was performed on the mRNA expression, and a heatmap was generated (Figure 2I).

qPCR Validation

To confirm the differential expression identified in our RNA-seq experiments, we used qPCR and analyzed 62 differentially expressed transcripts that were randomly selected: 26 circRNAs, 9 miRNAs, and 27 mRNAs. All selected transcripts were detected in the brain of 2- to 9-month-old APP/PS1 and wild-type mice, and exhibited statistically significant differential expression (Figures 3, 4, and 5). In summary, the qPCR results were highly consistent with the RNA-seq data.

Construction of circRNA-Associated ceRNA Networks

According to the ceRNA hypothesis, RNA transcripts effectively communicate with one another, and ceRNAs can compete for the same miRNA response elements (MREs) in regulatory networks. Here, we used our RNA-seq data to identify, for the first time, ceRNA networks in the APP/PS1 brain. We divided the differentially expressed transcripts (circRNAs, miRNAs, and mRNAs) into three groups: (1) 6yes9no, differential expression at 6 months, but not at 9 months of age; (2) 6no9yes, no differential expression at 6 months but differential expression at 9 months; and (3) 6yes9yes, differential expression at both 6 and 9 months. The transcripts of the 6yes9no, 6no9yes, and 6yes9yes groups could, respectively, play roles in AD pathogenesis, participate in AD development, and contribute to the disease at all stages (Figure 6A).

We selected 171 circRNAs and 1,104 mRNAs that were differentially expressed and shared common MRE binding sites (31 significantly dysregulated miRNAs) for inclusion in the 6yes9no group (Tables S7 and S8). A total of 136, 1,078, and 47 significantly dysregulated circRNAs, mRNAs, and miRNAs, respectively, were included in the 6no9yes group (Tables S9 and S10), and 3 circRNAs, 318 mRNAs, and 1 miRNA were included in the 6yes9yes group (Table S11). The ceRNA networks included both positive and negative regulation

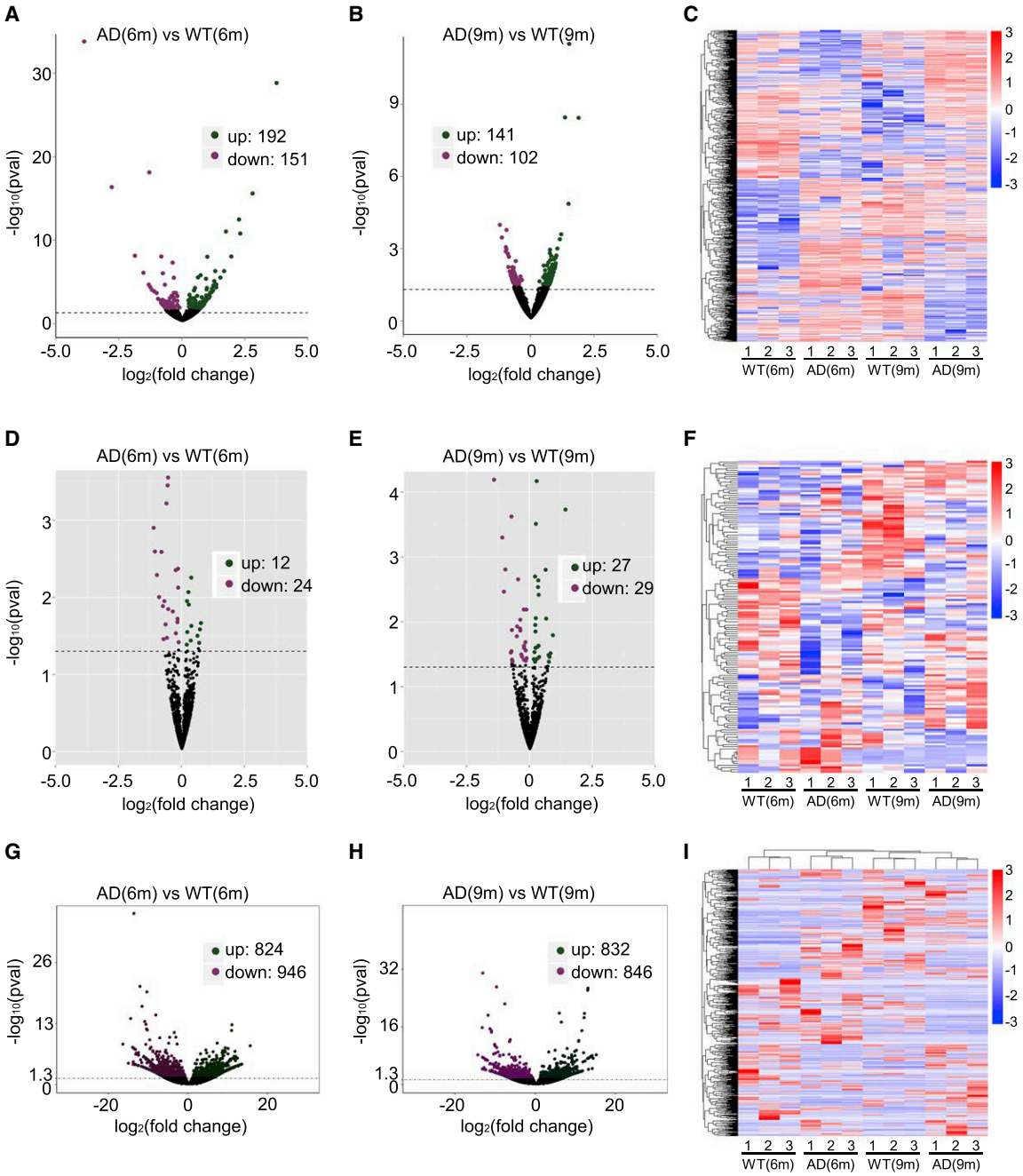
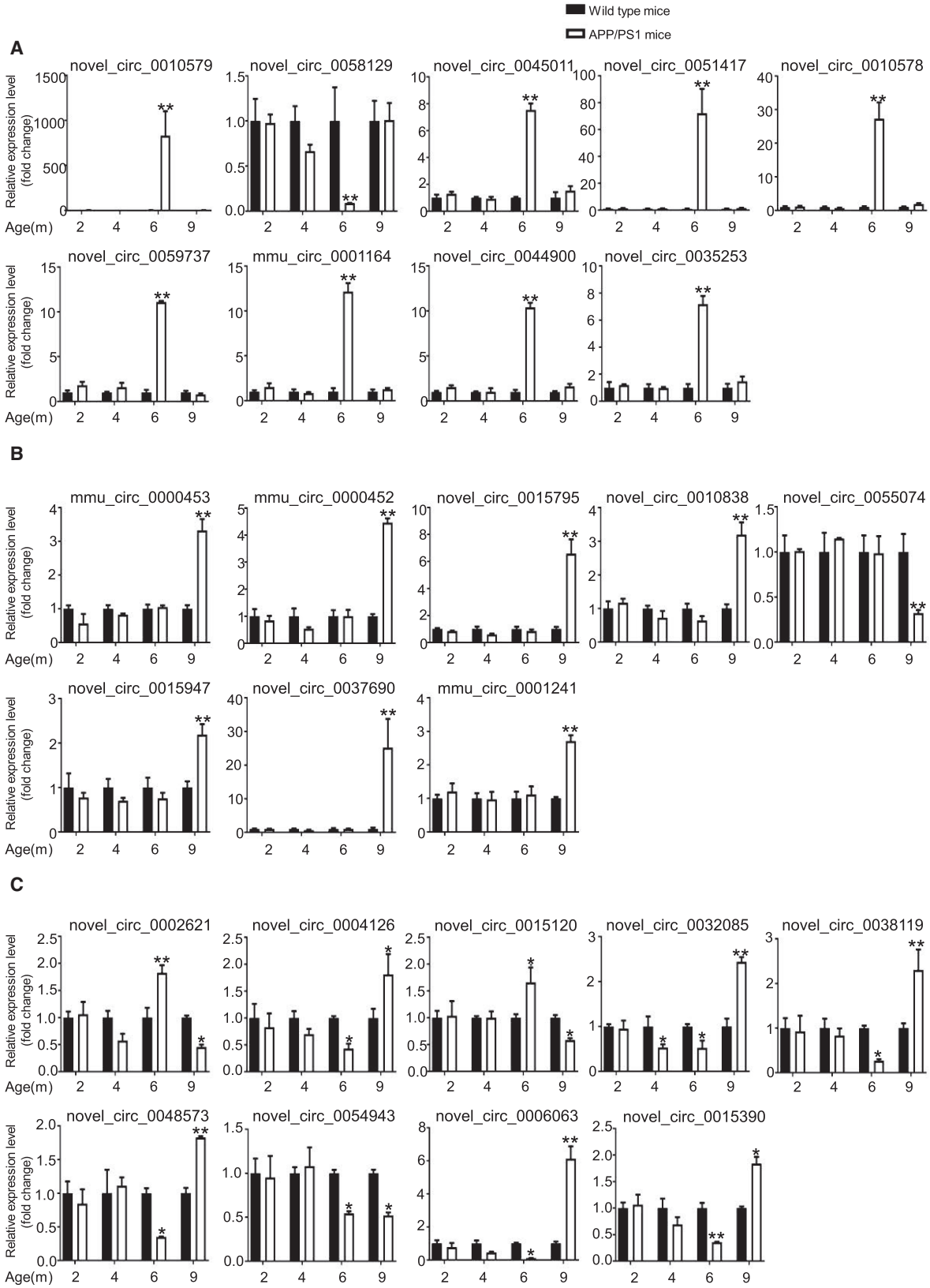


Figure 2. Expression Profiles of Distinct RNAs

(A–C) Expression profiles of circRNAs. (A and B) In the volcano plots, purple, green, and black points represent circRNAs that were downregulated, upregulated, and not significantly different in APP/PS1 mice relative to wild-type control mice at 6 and 9 months, respectively. x axis: log2 ratio of circRNA expression levels between AD and wild-type. y axis: false discovery rate values (−log10 transformed) of circRNAs. (C) Cluster analysis of expression of circRNAs. Red and blue: increased and decreased expression, respectively. (D–F) Expression profiles of miRNAs. (D and E) In the volcano plots, purple, green, and black points represent miRNAs that were downregulated, upregulated, and not significantly different in APP/PS1 mice relative to wild-type control mice at 6 and 9 months, respectively. x axis: log2 ratio of miRNA expression levels between AD and wild-type. y axis: false discovery rate values (−log10 transformed) of miRNAs. (F) Cluster analysis of expression of miRNAs. Red and blue: increased and decreased expression, respectively. (G–I) Expression profiles of mRNAs. (G and H) In the volcano plots, purple, green, and black points represent mRNAs that were downregulated, upregulated, and not significantly different in APP/PS1 mice relative to wild-type control mice at 6 and 9 months, respectively. x axis: log2 ratio of mRNA expression levels between AD and wild-type. y axis: false discovery rate values (−log10 transformed) of mRNAs. (I) Cluster analysis of expression of mRNAs. Red and blue: increased and decreased expression, respectively.



(legend on next page)

(Figures 6 and 7); Figure 6 shows the circRNAs, miRNAs, and mRNAs that were increased, decreased, and increased in APP/PS1 mice, respectively, and Figure 7 shows the circRNAs, miRNAs, and mRNAs that were decreased, increased, and decreased in APP/PS1 mice, respectively. These RNA interactions might be critical in AD pathogenesis.

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Pathway Analyses

A circRNA-associated ceRNA network can alter the regulation of the related mRNA-encoding genes. Gene Ontology (GO) analyses were performed on the genes in the networks identified here, and several GO terms were found to be significantly enriched (Tables S12, S13, and S14). The top highly enriched GO terms of biological process (BP), cellular component (CC), and molecular function (MF) are shown in Figure 8. The top terms were cytoskeleton (GO: 0005856), postsynaptic density (GO: 0014069), cell-cell adherens junction (GO: 0005913), and dendrite (GO: 0030425). Several cognition-associated terms were also observed, such as axon (GO: 0030424), synapse (GO: 0045202), postsynaptic density (GO: 0014069), intracellular signal transduction (GO: 0035556), and neuron projection (GO: 0043005). The enriched GO terms for the 6yes9no group differed from those for the 6no9yes group, which suggests changes in the functional genes during disease progression. In summary, circRNA-associated ceRNA networks might participate in the pathological progression of AD at distinct stages and through different mechanisms. The functions of the key genes in AD can now be studied, and the regulatory mechanisms in the ceRNA network can be investigated.

Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was conducted to determine the signaling cascades related to the identified genes; by using $p < 0.05$ as the threshold value, the following significantly enriched pathways were identified (Figure 9; Tables S12, S13, and S14): neuroactive ligand-receptor interaction, AMP-activated protein kinase (AMPK) signaling, long-term potentiation, Hippo signaling, glutamatergic synapse, phosphatidylinositol 3-kinase (PI3K)-Akt signaling, insulin secretion, focal adhesion, and axon guidance.

Association Study

We selectively analyzed data in the case of circRNAs, miRNAs, and their target genes that showed significant differential expression between APP/PS1 and wild-type mice (corrected $p < 0.05$); moreover, we used examples of the selected circRNAs, miRNAs, and their target genes that showed enrichment in the mouse brain to a certain order of magnitude, as well as those that were found to be associated with AD. We performed analyses under these conditions to investigate the most likely relationships between circRNA-associated ceRNA networks and AD. For example, mmu_circ_0000433, mmu_circ_0001473, novel_circ_0019965, novel_circ_0021924, novel_circ_0028455, novel_circ_0051361, and novel_circ_0058143 were identified as ceRNAs of mmu-miR-122b-3p, which targets *Cntnap2*. The expression of *Cntnap2* was higher in AD mice than in wild-type mice. *Cntnap2* influences the development of neural systems critical for learning and cross-modal integration, and disruption of this function could be associated with delayed learning.[33,34] Moreover, mmu_circ_0044900 was found to be a ceRNA of mmu-miR-449a-5p, mmu-miR-467a-3p, mmu-miR-540-3p, and mmu-miR-669f-3p, which target *Creb*. The p38-CREB-ICER molecules have been identified as key components of a negative-feedback mechanism necessary to regulate inflammation.[35] The ceRNA mmu_circ_0044900 inhibits miRNA functions by acting as a sponge for miRNAs. Our analysis also revealed that miRNAs can act directly on target genes. For example, mmu-miR-219b-5p, mmu-miR-350-5p, mmu-miR-450b-5p, and mmu-miR-9b-5p target *Sorbs2*, which influences dendritic development and memory.[36,37] Additional results are listed in Tables S7, S8, S9, S10, and S11. We predict that the identified circRNA-associated ceRNA networks are potentially involved in the regulation of AD.

DISCUSSION

AD is a fatal neurodegenerative disease that shows progressive development. Because of the complex pathogenesis of AD and the failure of drugs to cross the blood-brain barrier, effective diagnostic and therapeutic approaches for AD are lacking. Thus, several studies on AD have recently focused on the epigenetic regulation of AD pathogenesis to identify potential targets for therapy. For example, dynamic changes of DNA methylation and lncRNAs in the brain have been reported to contribute to AD.[38,39] However, the roles of circRNAs in AD have remained mostly unknown.

circRNAs lack 5′ caps and 3′ polyadenylated tails, and were therefore not detected in classical polyadenylated transcriptome studies. However, over the past few decades, the development of high-throughput sequencing and improved biochemical and computational biology methods have led to the identification of several circRNAs in various tissues and cells.[40] These circRNAs have been implicated in myriad human diseases, including AD.[18,41] Furthermore, circRNAs have been used as diagnostic and prognostic biomarkers because of the high exonuclease-digestion resistance conferred by their covalently closed loop structure.[42,43] For example, CDR1as[44] and circMTO1[45–47] have been reported as key players in the regulation of pathophysiological processes, which makes them targets for clinical diagnosis and treatment. However, our identification of circRNAs here is incomplete, and the specific roles of circRNAs in AD are largely unknown.

To the best of our knowledge, this is the first comprehensive high-throughput sequencing analysis of circRNA, miRNA, and mRNA

Figure 3. Validation of circRNA Expression by Using qPCR

The identified differentially expressed transcripts (circRNAs, miRNAs, and mRNAs) were divided into three groups: 6yes9no, differential expression at 6 months, but not at 9 months; 6no9yes, no differential expression at 6 months but differential expression at 9 months; and 6yes9yes, differential expression at both 6 and 9 months. (A–C) 6yes9no group (A), 6no9yes group (B), and 6yes9yes group (C). circRNA expression was quantified relative to *Gapdh* expression level by using the comparative cycle threshold (ΔCT) method. Data are presented as means ± SD (n = 3; *$p < 0.05$, **$p < 0.01$).

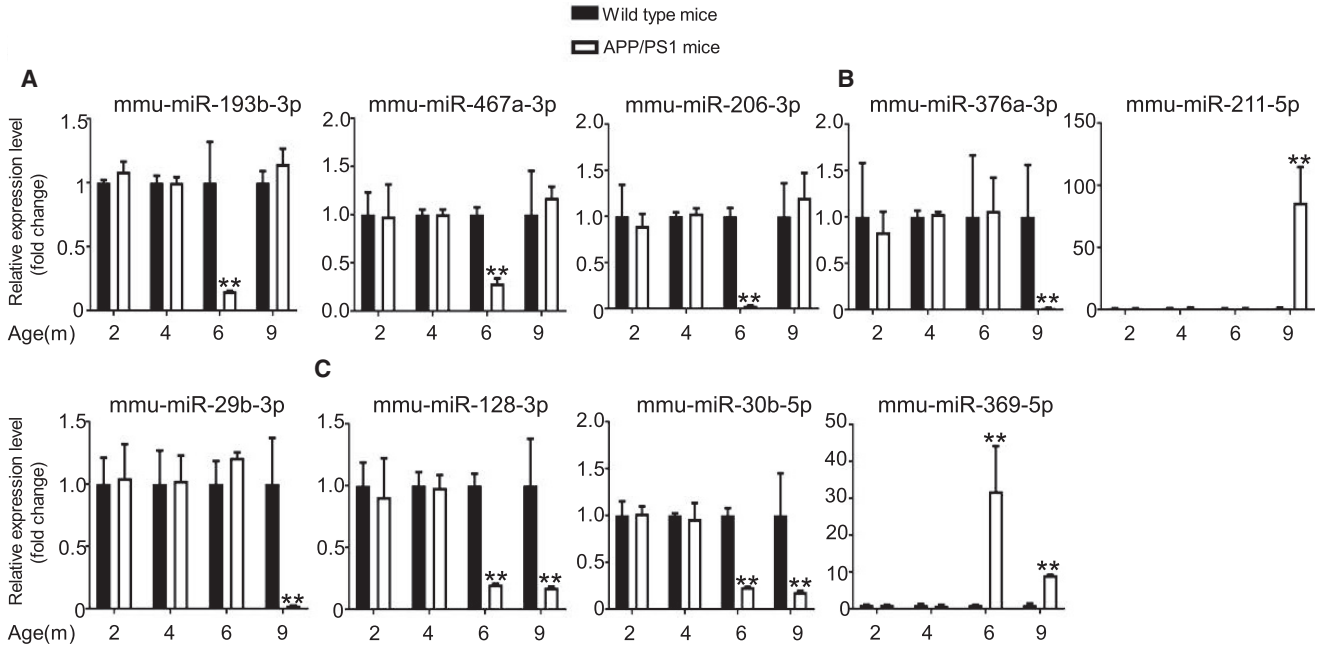


Figure 4. Validation of miRNA Expression by Using qPCR

(A–C) 6yes9no group (A), 6no9yes group (B), and 6yes9yes group (C). miRNA expression was quantified relative to *U6* expression level by using the comparative cycle threshold (ΔCT) method. Data are presented as means ± SD (n = 3; *p < 0.05, **p < 0.01).

expression profiles in the APP/PS1 mouse model of AD. At threshold values here of fold-change ≥ 2.0 and p < 0.05, dysregulated circRNAs, miRNAs, and mRNAs showed significant differential expression between AD and control groups. We consider these transcripts to be associated with the pathogenesis of AD. For instance, RTN4 interacts with BACE1 and inhibits its ability to produce Aβ[48]; ALS2 plays a critical role in neuronal development and degeneration[49]; and PDLIM5 is a homolog of AD-associated neuronal thread protein (AD7c-NTP), which is overexpressed in the early stage of AD, and overexpression of AD7c-NTP gene has been reported to cause neuritic sprouting and cell death.[50] The miRNA mmu-miR-124 is involved in the transformation of neuronal progenitor or non-neuronal cells into neuron-like cells,[51] and mmu-miR-134 regulates the development of cortical neurons.[52] The miRNA mmu-miR-107 is markedly downregulated in the early stages of AD, and the miRNA binds to the 3′ UTR of *BACE1* and increases *BACE1* mRNA levels, which indicates that mmu-miR-107 regulates the progression of BACE1-accelerated disease.[53] The level of mmu-miR-29 is also substantially reduced in AD patients, and mmu-miR-29 also regulates BACE1 expression.[54,55] Our qPCR validation study confirmed the profiles revealed by the high-throughput sequencing data, thus indicating the reliability of the sequencing data.

In this study, we used RNA-seq to systematically analyze circRNA, miRNA, and mRNA profiles in the brain of 6- and 9-month-old APP/PS1 mice. The transcripts of the 6yes9no group might play a crucial role in AD pathogenesis, and the stability and specific expression of these circRNAs could make them optimal biomarker candidates for AD; conversely, the transcripts in the 6no9yes group might function in AD development. Notably, transcripts in the 6yes9yes group are likely involved in the disease at all stages, and targeting these transcripts could facilitate the development of circRNA-based diagnostic tools and therapeutic strategies for AD. Overall, circRNA and miRNA molecules might act as key regulators of diverse aspects of AD. circRNAs and protein-coding mRNAs function as ceRNAs and super-sponges to regulate the expression of mRNA. Based on this theory, we obtained the miRNA-mRNA and miRNA-circRNA interaction data predicted by the tool miRanda and constructed a DEcircRNA-DEmiRNA-DEmRNA triple network for APP/PS1 and wild-type mouse brain. The selected circRNA-associated ceRNA networks could offer new insights into AD and thus suggest novel treatments for the disease.

We performed GO enrichment and KEGG analysis of the genes in the ceRNA networks and identified enriched terms that are relevant to the pathological process of AD, including synapse (GO: 0045202), cytoskeleton (GO: 0005856), postsynaptic density (GO: 0014069), cell-cell adherens junction (GO: 0005913), dendrite (GO: 0030425), axon (GO: 0030424), and neuron projection (GO: 0043005), as well as various pathways, including cAMP signaling, mitogen-activated protein kinase (MAPK) signaling, insulin secretion, Hippo signaling, adherens junction, focal adhesion, dopaminergic synapse, and PI3K-Akt signaling pathways. Analysis of the data revealed networks that participate in AD. One of these networks involves the gene *Scube2*. The ceRNAs mmu_circ_0000452, mmu_circ_0000453, novel_circ_0010838, novel_circ_0011428, novel_circ_0033961, and novel_circ_0037760 might

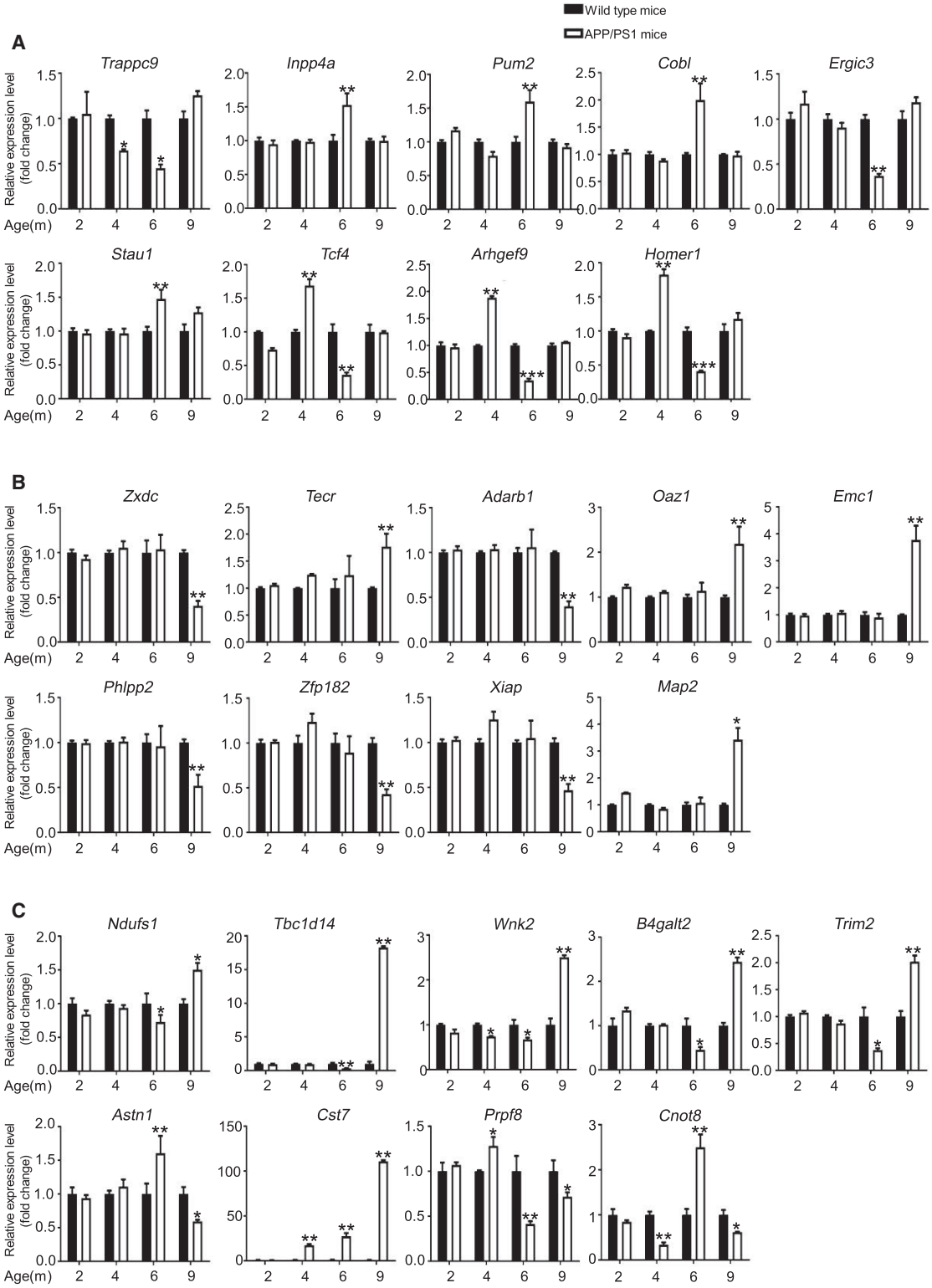


Figure 5. Validation of mRNA Expression by Using qPCR

(A–C) 6yes9no group (A), 6no9yes group (B), and 6yes9yes group (C). mRNA expression was quantified relative to *Gapdh* expression level by using the comparative cycle threshold (ΔCT) method. Data are presented as means ± SD (n = 3; *p < 0.05, **p < 0.01).

Figure 6. circRNA-Associated ceRNA Networks in APP/PS1 Mice

ceRNA networks were constructed based on identified circRNA-miRNA and miRNA-mRNA interactions. The networks include increased circRNAs in APP/PS1 mice, decreased miRNAs in APP/PS1 mice, and increased mRNAs in APP/PS1 mice. (A–D) Grouping (A), 6yes9no group (B), 6no9yes group (C), and 6yes9yes group (D).

bind to mmu-miR-466b-5p, which targets *Scube2*, a newly identified mammalian epidermal growth factor-related gene that plays a key role in mouse neural development and is critical for AD progression.[56] The circRNA-associated ceRNA networks in AD are highly complicated and diverse. The APP/PS1 mouse model used in this study cannot represent the entire disease, which is mainly related to β-amyloid toxicity. Therefore, further investigation is required to understand the regulation of these networks in AD. In a recent study, circRNAs have been found to be associated with the pathogenesis in AD patients. The expression of circRNA might have strong predictive ability for human AD case status, indicating that the mechanism of circRNA in AD development needs to be further elucidated.[57]

In conclusion, we elucidated the brain circRNA-associated ceRNA profiles of APP/PS1 and wild-type mice by using deep RNA-seq analysis. Our findings further expand the current knowledge regarding the biology of ceRNAs and their regulatory roles in AD pathogenesis. These newly identified networks reveal potential biomarkers or therapeutic targets for AD, and our results should serve as a valuable resource for the clinical diagnosis, treatment, and prevention of AD.

MATERIALS AND METHODS

Tissue Preparation

Wild-type and APP/PS1 mice (originally from The Jackson Laboratory; strain name: "B6.Cg-Tg(APPswe, PSEN1dE9)85Dbo/Mmjax[46]") were purchased from the Model Animal Research Center of Nanjing University. The mice were housed one per cage under standard specific conditions (25°C, 50% humidity, 12/12-h light/dark cycle, and pathogen-free environment). The mice were provided free access to standard diet until they met age requirements (6 and 9 months); then three male mice were randomly selected from each group, and their cerebral cortex was collected for RNA-seq. All animal experiments were performed in accordance with animal use protocols approved by the Committee for the Ethics of Animal Experiments, Shenzhen Peking University–The Hong Kong University of Science and Technology Medical Center (SPHMC) (protocol number 2011-004).

RNA Extraction and Qualification

Total RNA from each sample was isolated using TRIzol reagent (Invitrogen), according to the manufacturer's instructions, and then 1% agarose gels were used to monitor RNA degradation and contamination. RNA purity was assessed using a NanoPhotometer spectrophotometer (IMPLEN, CA, USA), RNA concentration was measured

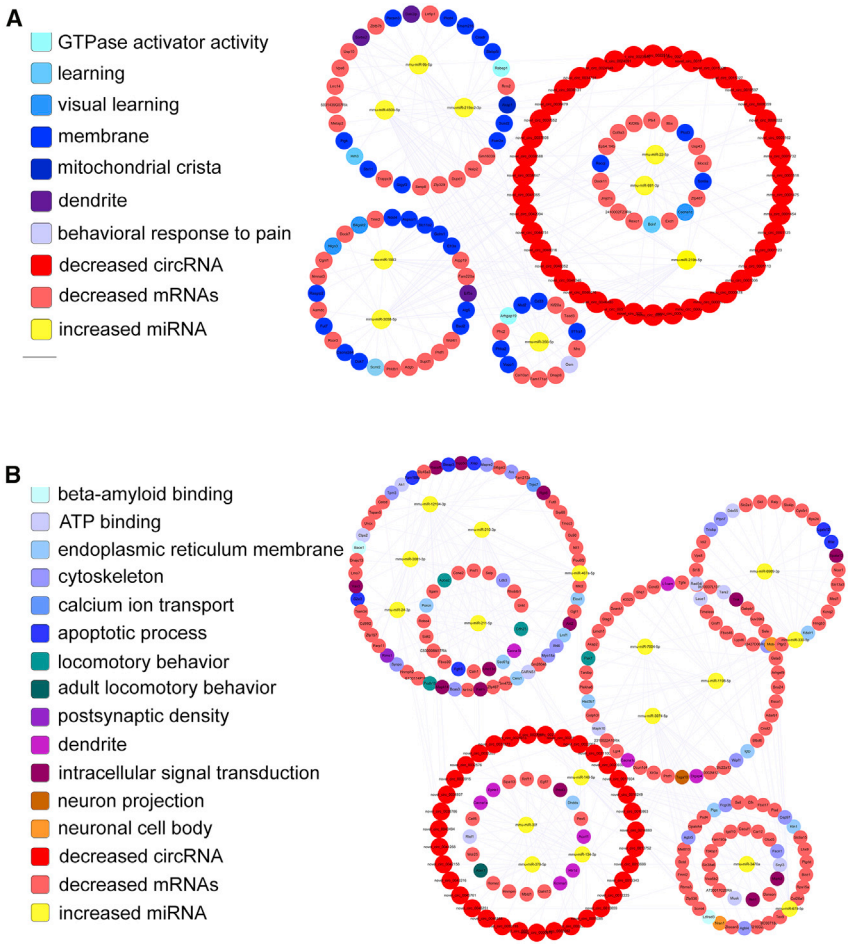


Figure 7. circRNA-Associated ceRNA Networks in APP/PS1 Mice

ceRNA networks were constructed based on identified circRNA-miRNA and miRNA-mRNA interactions. The networks include decreased circRNAs in APP/PS1 mice, increased miRNAs in APP/PS1 mice, and decreased mRNAs in APP/PS1 mice. (A and B) 6yes9no group (A) and 6no9yes group (B).

using a Qubit RNA Assay Kit in a Qubit 2.0 Fluorometer (Life Technologies, CA, USA), and RNA integrity was evaluated using the RNA Nano 6000 Assay Kit of a Bioanalyzer 2100 System (Agilent Technologies, CA, USA).

RNA-Seq

Details of the mRNA-seq, miRNA-seq, and circRNA-seq methods are described in Document S1.

Expression Analysis

We calculated the FPKMs values of transcripts by using Cuffdiff (v.2.1.1) to evaluate the expression levels of protein-coding genes in each sample.[58] The expression levels of miRNAs and circRNAs were estimated in terms of TPM values by following these criteria[58]: transcripts featuring p values <0.05 were regarded as being differentially expressed between APP/PS1 and wild-type mice: Normalized expression = (mapped reads)/(total reads) × 1,000,000.

ceRNA Network Analysis

The expression levels of circRNAs, miRNAs, and mRNAs differed significantly between APP/PS1 and wild-type mice, and the RNAs were therefore further analyzed. We searched the sequences of the circRNAs and mRNAs to identify potential MREs. We used miRanda (http://www.microrna.org/microrna/home.do) to predict miRNA-binding seed-sequence sites, and an overlap of the same miRNA-binding sites on both circRNAs and mRNAs indicated potential circRNA-miRNA-mRNA interaction.

GO Annotations and KEGG Pathway Analyses

The DAVID (https://david.ncifcrf.gov/summary.jsp) database was used to analyze circRNA-miRNA-enriched genes. GO and KEGG terms featuring p values <0.05 were considered significantly enriched.

circRNA-Associated ceRNA Network Construction

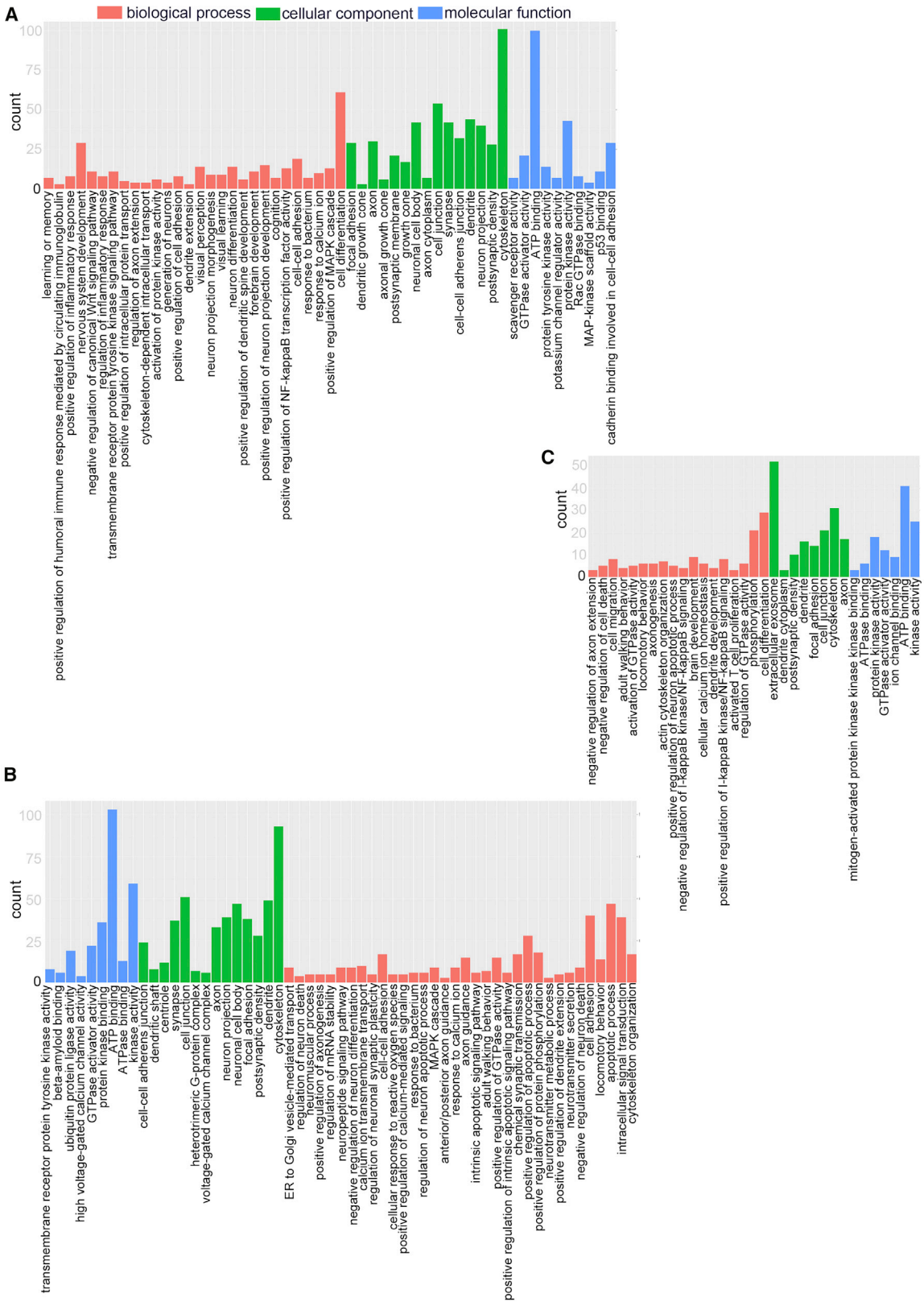
circRNA-associated ceRNA networks were constructed and visually displayed using Cytoscape software v.3.5.0 (San Diego, CA, USA) based on the results obtained from analyzing high-throughput sequencing data, as described above. In the figures, distinct shapes and colors represent different RNA types, GO terms, and regulatory relationships.

Quantitative Real-Time PCR Validation

Total RNA was extracted using TRIzol reagent (Sigma) according to the manufacturer's protocol. RNA quantity was measured using a NanoDrop 2000 (Thermo Fisher Scientific). qRT-PCR was performed using the GoScript Reverse Transcription System (Promega) in a C1000 Thermal Cycler (Bio-Rad). The glyceraldehyde-3-phosphate dehydrogenase gene (*Gapdh*) was used as an internal control. Relative gene-expression levels were calculated using the $2^{-\Delta CT}$ method.

Statistical Analysis

Two normally distributed groups were compared using t tests. Parameters for the high-throughput sequencing-related data were calculated, and statistical computing was performed using R software. All data are expressed as means ± SD; $p < 0.05$ was considered statistically significant.



(legend on next page)

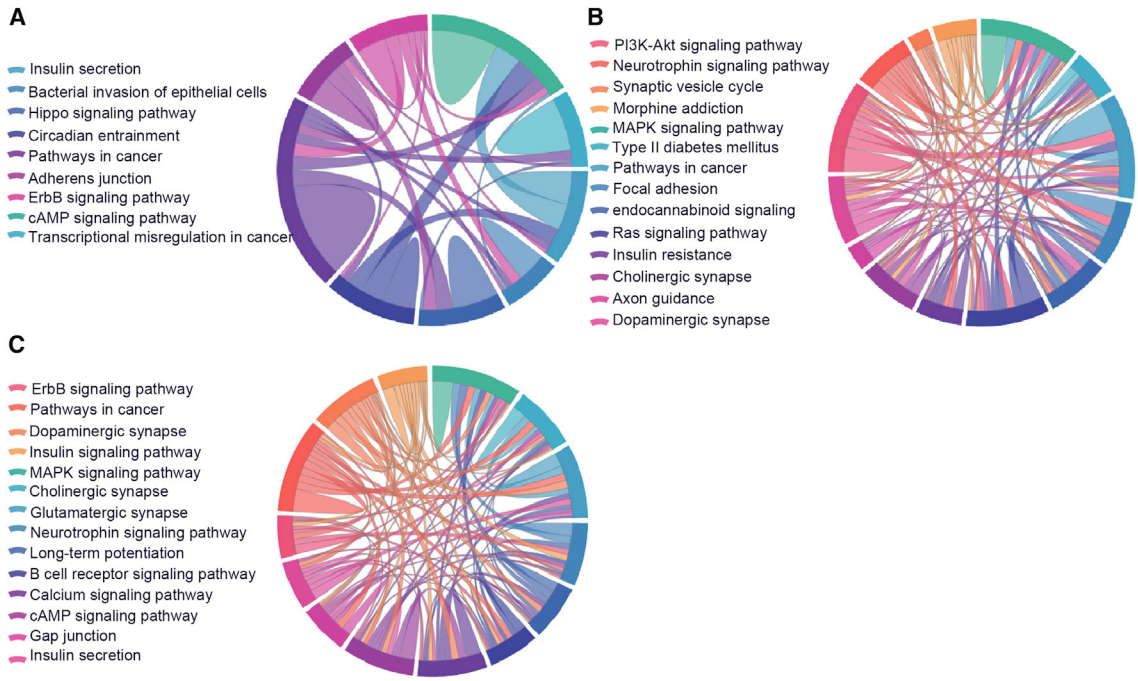


Figure 9. Significantly Enriched Kyoto Encyclopedia of Genes and Genomes (KEGG) and Reactome Pathways

circRNA-associated ceRNA network genes participate in distinct aspects of AD pathology. (A–C) 6yes9no group (A), 6no9yes group (B), and 6yes9yes group (C). Significantly enriched KEGG pathways featured p values <0.05. Each line represents a gene, and the number of lines indicates the genes enriched. KEGG analysis was conducted using the DAVID (https://david.ncifcrf.gov/summary.jsp) database.

Data Access

All raw and processed sequencing data generated in this study have been submitted to the NCBI GEO (https://www.ncbi.nlm.nih.gov/geo/) database under accession number GEO: GSE132177.

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.omtn.2019.10.030.

AUTHOR CONTRIBUTIONS

J.W., W.Z., and N.M. designed the study. N.M., J.P., X.Y., and B.Y. participated in the animal experiments including tissue collection and RNA extraction. N.M. performed the experiments and analyzed the data. N.M., J.W., and W.Z. analyzed the results and wrote the manuscript. All authors read and approved the final manuscript.

CONFLICTS OF INTEREST

The authors declare no competing interests.

ACKNOWLEDGMENTS

We would like to thank the Shenzhen Biomedical Research Support Platform and the Shenzhen Molecular Diagnostic Platform of Dermatology for technical help. We thank Dr. Madhavan Raghavan for critical reading of the manuscript. This work was supported by the National Key R&D Program of China (grant 2016YFA0501903); National Natural Scientific Foundation of China (grants 81571043, 81673053, and 81802860); Natural Scientific Foundation of Guangdong Province (grant 2016A030312016); and Shenzhen Basic Research (grants JCYJ20160229153100269, JCYJ20170411090739316, JCYJ20170815153617033, JCYJ20180507182657867, JCYJ20170306161450254, and JCYJ20170306161807726).

REFERENCES

1. Zádori, D., Veres, G., Szalárdy, L., Klivényi, P., and Vécsei, L. (2018). Alzheimer's Disease: Recent Concepts on the Relation of Mitochondrial Disturbances, Excitotoxicity, Neuroinflammation, and Kynurenines. J. Alzheimers Dis. *62*, 523–547.

Figure 8. Gene Ontology (GO) Enrichment Annotations of Pathological Progression of AD: Biological Process, Cellular Component, and Molecular Function

The top terms were synapse (GO: 0045202), cytoskeleton (GO: 0005856), postsynaptic density (GO: 0014069), cell-cell adherens junction (GO: 0005913), dendrite (GO: 0030425), axon (GO: 0030424), and neuron projection (GO: 0043005). (A–C) 6yes9no group (A), 6no9yes group (B), and 6yes9yes group (C). Significantly enriched GO pathways featured p values <0.05. GO analysis was conducted using the DAVID (https://david.ncifcrf.gov/summary.jsp) database.

2. Alzheimer's Association (2018). 2018 Alzheimer's disease facts and figures. Alzheimers Dement. *14*, 367–429.
3. Tan, L., Yu, J.T., Hu, N., and Tan, L. (2013). Non-coding RNAs in Alzheimer's disease. Mol. Neurobiol. *47*, 382–393.
4. Millan, M.J. (2017). Linking deregulation of non-coding RNA to the core pathophysiology of Alzheimer's disease: An integrative review. Prog. Neurobiol. *156*, 1–68.
5. Chen, L.L. (2016). The biogenesis and emerging roles of circular RNAs. Nat. Rev. Mol. Cell Biol. *17*, 205–211.
6. Memczak, S., Jens, M., Elefsinioti, A., Torti, F., Krueger, J., Rybak, A., Maier, L., Mackowiak, S.D., Gregersen, L.H., Munschauer, M., et al. (2013). Circular RNAs are a large class of animal RNAs with regulatory potency. Nature *495*, 333–338.
7. Rybak-Wolf, A., Stottmeister, C., Glažar, P., Jens, M., Pino, N., Giusti, S., Hanan, M., Behm, M., Bartok, O., Ashwal-Fluss, R., et al. (2015). Circular RNAs in the Mammalian Brain Are Highly Abundant, Conserved, and Dynamically Expressed. Mol. Cell *58*, 870–885.
8. Westholm, J.O., Miura, P., Olson, S., Shenker, S., Joseph, B., Sanfilippo, P., Celniker, S.E., Graveley, B.R., and Lai, E.C. (2014). Genome-wide analysis of drosophila circular RNAs reveals their structural and sequence properties and age-dependent neural accumulation. Cell Rep. *9*, 1966–1980.
9. Gruner, H., Cortés-López, M., Cooper, D.A., Bauer, M., and Miura, P. (2016). CircRNA accumulation in the aging mouse brain. Sci. Rep. *6*, 38907.
10. Hansen, T.B., Jensen, T.I., Clausen, B.H., Bramsen, J.B., Finsen, B., Damgaard, C.K., and Kjems, J. (2013). Natural RNA circles function as efficient microRNA sponges. Nature *495*, 384–388.
11. Han, D., Li, J., Wang, H., Su, X., Hou, J., Gu, Y., Qian, C., Lin, Y., Liu, X., Huang, M., et al. (2017). Circular RNA circMTO1 acts as the sponge of microRNA-9 to suppress hepatocellular carcinoma progression. Hepatology *66*, 1151–1164.
12. Zheng, J., Liu, X., Xue, Y., Gong, W., Ma, J., Xi, Z., Que, Z., and Liu, Y. (2017). TTBK2 circular RNA promotes glioma malignancy by regulating miR-217/HNF1β/Derlin-1 pathway. J. Hematol. Oncol. *10*, 52.
13. Huang, M., Zhong, Z., Lv, M., Shu, J., Tian, Q., and Chen, J. (2016). Comprehensive analysis of differentially expressed profiles of lncRNAs and circRNAs with associated co-expression and ceRNA networks in bladder carcinoma. Oncotarget *7*, 47186–47200.
14. Fan, X., Weng, X., Zhao, Y., Chen, W., Gan, T., and Xu, D. (2017). Circular RNAs in Cardiovascular Disease: An Overview. Biomed. Res. Int. *2017*, 5135781.
15. Peng, L., Chen, G., Zhu, Z., Shen, Z., Du, C., Zang, R., Su, Y., Xie, H., Li, H., Xu, X., et al. (2017). Circular RNA ZNF609 functions as a competitive endogenous RNA to regulate AKT3 expression by sponging miR-150-5p in Hirschsprung's disease. Oncotarget *8*, 808–818.
16. Bingol, B., and Sheng, M. (2011). Deconstruction for reconstruction: the role of proteolysis in neural plasticity and disease. Neuron *69*, 22–32.
17. Lonskaya, I., Shekoyan, A.R., Hebron, M.L., Desforges, N., Algarzae, N.K., and Moussa, C.E. (2013). Diminished parkin solubility and co-localization with intraneuronal amyloid-β are associated with autophagic defects in Alzheimer's disease. J. Alzheimers Dis. *33*, 231–247.
18. Lukiw, W.J. (2013). Circular RNA (circRNA) in Alzheimer's disease (AD). Front. Genet. *4*, 307.
19. Tatro, E.T., Risbrough, V., Soontornniyomkij, B., Young, J., Shumaker-Armstrong, S., Jeste, D.V., and Achim, C.L. (2013). Short-term recognition memory correlates with regional CNS expression of microRNA-138 in mice. Am. J. Geriatr. Psychiatry *21*, 461–473.
20. Schröder, J., Ansaloni, S., Schilling, M., Liu, T., Radke, J., Jaedicke, M., Schjeide, B.M., Mashychev, A., Tegeler, C., Radbruch, H., et al. (2014). MicroRNA-138 is a potential regulator of memory performance in humans. Front. Hum. Neurosci. *8*, 501.
21. Jankowsky, J.L., Fadale, D.J., Anderson, J., Xu, G.M., Gonzales, V., Jenkins, N.A., Copeland, N.G., Lee, M.K., Younkin, L.H., Wagner, S.L., et al. (2004). Mutant presenilins specifically elevate the levels of the 42 residue beta-amyloid peptide in vivo: evidence for augmentation of a 42-specific gamma secretase. Hum. Mol. Genet. *13*, 159–170.
22. Wang, Z., Gerstein, M., and Snyder, M. (2009). RNA-Seq: a revolutionary tool for transcriptomics. Nat. Rev. Genet. *10*, 57–63.
23. Marguerat, S., and Bähler, J. (2010). RNA-seq: from technology to biology. Cell. Mol. Life Sci. *67*, 569–579.
24. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. Nat. Methods *9*, 357–359.
25. Gao, Y., Zhang, J., and Zhao, F. (2018). Circular RNA identification based on multiple seed matching. Brief Bioinform. *19*, 803–810.
26. Wen, M., Shen, Y., Shi, S., and Tang, T. (2012). miREvo: an integrative microRNA evolutionary analysis platform for next-generation sequencing experiments. BMC Bioinformatics *13*, 140.
27. Friedländer, M.R., Mackowiak, S.D., Li, N., Chen, W., and Rajewsky, N. (2012). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res. *40*, 37–52.
28. Kucherov, G., Salikhov, K., and Tsur, D. (2014). Approximate string matching using a bidirectional index. arXiv, arXiv:1310.1440v2, https://arxiv.org/abs/1310.1440v2.
29. Sun, L., Luo, H., Bu, D., Zhao, G., Yu, K., Zhang, C., Liu, Y., Chen, R., and Zhao, Y. (2013). Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. Nucleic Acids Res. *41*, e166.
30. Kang, Y.J., Yang, D.C., Kong, L., Hou, M., Meng, Y.Q., Wei, L., and Gao, G. (2017). CPC2: a fast and accurate coding potential calculator based on sequence intrinsic features. Nucleic Acids Res. *45* (W1), W12–W16.
31. Mistry, J., Bateman, A., and Finn, R.D. (2007). Predicting active site residue annotations in the Pfam database. BMC Bioinformatics *8*, 298.
32. Lin, M.F., Jungreis, I., and Kellis, M. (2011). PhyloCSF: a comparative genomics method to distinguish protein coding and non-coding regions. Bioinformatics *27*, i275–i282.
33. Rendall, A.R., Truong, D.T., and Fitch, R.H. (2016). Learning delays in a mouse model of Autism Spectrum Disorder. Behav. Brain Res. *303*, 201–207.
34. Canali, G., Garcia, M., Hivert, B., Pinatel, D., Goullancourt, A., Oguievetskaia, K., Saint-Martin, M., Girault, J.A., Faivre-Sarrailh, C., and Goutebroze, L. (2018). Genetic variants in autism-related CNTNAP2 impair axonal growth of cortical neurons. Hum. Mol. Genet. *27*, 1941–1954.
35. Lv, S., Li, J., Qiu, X., Li, W., Zhang, C., Zhang, Z.N., and Luan, B. (2017). A negative feedback loop of ICER and NF-κB regulates TLR signaling in innate immune responses. Cell Death Differ. *24*, 492–499.
36. Zhang, Q., Gao, X., Li, C., Feliciano, C., Wang, D., Zhou, D., Mei, Y., Monteiro, P., Anand, M., Itohara, S., et al. (2016). Impaired Dendritic Development and Memory in Sorbs2 Knock-Out Mice. J. Neurosci. *36*, 2247–2260.
37. Lee, S.E., and Chang, S. (2016). nArgBP2 as a hub molecule in the etiology of various neuropsychiatric disorders. BMB Rep. *49*, 457–458.
38. Zhang, S., Qin, C., Cao, G., Xin, W., Feng, C., and Zhang, W. (2016). Systematic Analysis of Long Noncoding RNAs in the Senescence-accelerated Mouse Prone 8 Brain Using RNA Sequencing. Mol. Ther. Nucleic Acids *5*, e343.
39. Zhang, S., Qin, C., Cao, G., Guo, L., Feng, C., and Zhang, W. (2017). Genome-wide analysis of DNA methylation profiles in a senescence-accelerated mouse prone 8 brain using whole-genome bisulfite sequencing. Bioinformatics *33*, 1591–1595.
40. Chen, L.L., and Yang, L. (2015). Regulation of circRNA biogenesis. RNA Biol. *12*, 381–388.
41. Zhao, Y., Alexandrov, P.N., Jaber, V., and Lukiw, W.J. (2016). Deficiency in the Ubiquitin Conjugating Enzyme UBE2A in Alzheimer's Disease (AD) is Linked to Deficits in a Natural Circular miRNA-7 Sponge (circRNA; ciRS-7). Genes (Basel) *7*, 116.
42. Li, P., Chen, S., Chen, H., Mo, X., Li, T., Shao, Y., Xiao, B., and Guo, J. (2015). Using circular RNA as a novel type of biomarker in the screening of gastric cancer. Clin. Chim. Acta *444*, 132–136.
43. Jiang, X.M., Li, Z.L., Li, J.L., Xu, Y., Leng, K.M., Cui, Y.F., and Sun, D.J. (2018). A novel prognostic biomarker for cholangiocarcinoma: circRNA Cdr1as. Eur. Rev. Med. Pharmacol. Sci. *22*, 365–371.
44. Yang, X., Xiong, Q., Wu, Y., Li, S., and Ge, F. (2017). Quantitative Proteomics Reveals the Regulatory Networks of Circular RNA CDR1as in Hepatocellular Carcinoma Cells. J. Proteome Res. *16*, 3891–3902.

45. Han, D., Li, J., Wang, H., Su, X., Hou, J., Gu, Y., Qian, C., Lin, Y., Liu, X., Huang, M., et al. (2017). Circular RNA circMTO1 acts as the sponge of microRNA-9 to suppress hepatocellular carcinoma progression. Hepatology *66*, 1151–1164.

46. Rao, J., Cheng, X., Zhu, H., Wang, L., and Liu, L. (2018). Retraction: Circular RNA-0007874 (circMTO1) reverses chemoresistance to temozolomide by acting as a sponge of microRNA-630 in glioblastoma. Cell Biol. Int. Published online December 11, 2018. https://doi.org/10.1002/cbin.11080.

47. Li, Y., Wan, B., Liu, L., Zhou, L., and Zeng, Q. (2019). Circular RNA circMTO1 suppresses bladder cancer metastasis by sponging miR-221 and inhibiting epithelial-to-mesenchymal transition. Biochem. Biophys. Res. Commun. *508*, 991–996.

48. Murayama, K.S., Kametani, F., Saito, S., Kume, H., Akiyama, H., and Araki, W. (2006). Reticulons RTN3 and RTN4-B/C interact with BACE1 and inhibit its ability to produce amyloid beta-protein. Eur. J. Neurosci. *24*, 1237–1244.

49. Hadano, S., Kunita, R., Otomo, A., Suzuki-Utsunomiya, K., and Ikeda, J.E. (2007). Molecular and cellular function of ALS2/alsin: implication of membrane dynamics in neuronal development and degeneration. Neurochem. Int. *51*, 74–84.

50. Wu, M., Li, Y., Ji, C., Xu, J., Zheng, H., Zou, X., Gu, S., Lou, Y., Xie, Y., and Mao, Y. (2004). Cloning and identification of a novel human gene PDLIM5, a homolog of AD-associated neuronal thread protein (AD7c-NTP). DNA Seq. *15*, 144–147.

51. Makeyev, E.V., Zhang, J., Carrasco, M.A., and Maniatis, T. (2007). The MicroRNA miR-124 promotes neuronal differentiation by triggering brain-specific alternative pre-mRNA splicing. Mol. Cell *27*, 435–448.

52. Gaughwin, P., Ciesla, M., Yang, H., Lim, B., and Brundin, P. (2011). Stage-specific modulation of cortical neuronal development by Mmu-miR-134. Cereb. Cortex *21*, 1857–1869.

53. Nelson, P.T., and Wang, W.X. (2010). MiR-107 is reduced in Alzheimer's disease brain neocortex: validation study. J. Alzheimers Dis. *21*, 75–79.

54. Shioya, M., Obayashi, S., Tabunoki, H., Arima, K., Saito, Y., Ishida, T., and Satoh, J. (2010). Aberrant microRNA expression in the brains of neurodegenerative diseases: miR-29a decreased in Alzheimer disease brains targets neurone navigator 3. Neuropathol. Appl. Neurobiol. *36*, 320–330.

55. Roshan, R., Ghosh, T., Gadgil, M., and Pillai, B. (2012). Regulation of BACE1 by miR-29a/b in a cellular model of Spinocerebellar Ataxia 17. RNA Biol. *9*, 891–899.

56. Grimmond, S., Larder, R., Van Hateren, N., Siggers, P., Morse, S., Hacker, T., Arkell, R., and Greenfield, A. (2001). Expression of a novel mammalian epidermal growth factor-related gene during mouse neural development. Mech. Dev. *102*, 209–211.

57. Dube, U., Del-Aguila, J.L., Li, Z., Budde, J.P., Jiang, S., Hsu, S., Ibanez, L., Fernandez, M.V., Farias, F., Norton, J., et al. (2019). An atlas of cortical circular RNA expression in Alzheimer disease brains demonstrates clinical and pathological associations. Nat. Neurosci. *22*, 1903–1912.

58. Trapnell, C., Williams, B.A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M.J., Salzberg, S.L., Wold, B.J., and Pachter, L. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. *28*, 511–515.

OMTN, Volume *18*

Supplemental Information

Whole-Transcriptome Analysis of APP/PS1 Mouse Brain and Identification of circRNA-miRNA-mRNA Networks to Investigate AD Pathogenesis

Nana Ma, Jie Pan, Xiaoyang Ye, Bo Yu, Wei Zhang, and Jun Wan

CircRNA Methods

RNA isolation, quantification and qualification

RNA degradation and contamination was monitored on 1% agarose gels. RNA purity was checked using the Nano Photometer® spectrophotometer (IMPLEN, CA, USA). RNA concentration was measured using Qubit® RNA Assay Kit in Qubit® 2.0 Flurometer (Life Technologies, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Library preparation for circRNA sequencing

A total amount of 5 μg RNA per sample was used as input material for the RNA sample preparations. Firstly, ribosomal RNA was removed by Epicentre Ribo- zero™ rRNA Removal Kit (Epicentre, USA), and rRNA free residue was cleaned up by ethanol precipitation. Subsequently, the linear RNA was digested with 3 U of RNase R (Epicentre, USA) per μg of RNA. The sequencing libraries were generated by NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations. Briefly, fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer(5X). First strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNaseH-). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. In the reaction buffer, dNTPs with dTTP were replaced by dUTP. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends

of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 150~200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then 3 μl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37° C for 15 min followed by 5 min at 95°C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system polymerase, Universal PCR primers and Index (X) Primer. At last, products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system.

Clustering and sequencing

The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumia) according to the manufacturer's instructions. After cluster generation, the libraries were sequenced on an Illumina Hiseq 4000 platform and 150 bp paired- end reads were generated.

Data analysis

Quality control

Raw data (raw reads) of fastq format were firstly processed through in-house perlscripts. In this step, clean data (clean reads) were obtained by removing reads containing adapter, reads

on containing ploy-N and low quality reads from raw data. At the same time, Q20, Q30 and GC content of the clean data were calculated. All the downstream analyses were based on the clean data with high quality.

Mapping to the reference genome

Reference genome and gene model annotation files were downloaded from genome website directly. Index of the reference genome was built using bowtie2 v2.2.8 and paired-end clean reads were aligned to the reference genome using Bowtie (Langmead, B.et al).

circRNA identification

The circRNA were detected and identified using find_circ (Sebastian Memczak et al., 2013) and CIRI2 (Gao et al., 2017). Circos software was used to construct the circos figure.

Quantification of gene expression level

The raw counts were first normalized using TPM (Zhou et al., 2010)

Normalized expression level = (readCount*1,000,000)/libsize (libsize is the sum of circRNA readcount).

Differential expression analysis

Differential expression analysis of two conditions/groups was performed using the DESeq R package (1.10.1). DESeq provide statistical routines for determining differential expression in digital gene expression data using a model based on the negative binomial distribution. The

resulting P values were adjusted using the Benjamini and Hochberg's approach for controlling the false discovery rate. Genes with an adjusted P-value found by DESeq were assigned as differentially expressed.

MicroRNA targer site analysis

MicroRNA targer site in exons of circRNA loci were identified using miRanda (animal species).

CircRNA-miRNA-gene network analysis

Cytoscape software was used to construct the circRNA-miRNA-gene networks.

microRNA Methods

RNA isolation, quantification and qualification

RNA degradation and contamination was monitored on 1% agarose gels. RNA purity was checked using the Nano Photometer® spectrophotometer (IMPLEN, CA, USA). RNA concentration was measured using Qubit® RNA Assay Kit in Qubit® 2.0 Flurometer (Life Technologies, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Library preparation for Small RNA sequencing

A total amount of 3 μg total RNA per sample was used as input material for the small RNA library. Sequencing libraries were generated using NEBNext® Multiplex Small RNA Library Prep Set for Illumina® (NEB, USA.) following manufacturer's recommendations and index codes were added to attribute sequences to each sample. Briefly, NEB 3' SR Adaptor was directly, and specifically ligated to 3' end of miRNA, siRNA and piRNA. After the 3' ligation reaction, the SR RT Primer hybridized to the excess of 3' SR Adaptor (that remained free after the 3' ligation reaction) and transformed the single-stranded DNA adaptor into a double-stranded DNA molecule. This step is important to prevent adaptor-dimer formation, besides, dsDNAs are not substrates for ligation mediated by T4 RNA Ligase 1 and therefore do not ligate to the 5´ SR Adaptor in the subsequent ligation step. 5´ends adapter was ligated to 5´ends of miRNAs, siRNA and piRNA. Then first strand cDNA was synthesized using M-MuLV Reverse Transcriptase (RNase H–). PCR amplification was performed using LongAmp Taq 2X Master Mix, SR Primer for illumina and index (X) primer. PCR products were purified on a 8% polyacrylamide gel (100V, 80 min). DNA fragments corresponding to 140~160 bp (the length of small noncoding RNA plus the 3' and 5' adaptors) were recovered and dissolved in 8 μL elution buffer. At last, library quality was assessed on the Agilent Bioanalyzer 2100 system using DNA High Sensitivity Chips.

Clustering and sequencing

The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq SR Cluster Kit v3-cBot-HS (Illumia) according to the manufacturer's instructions. After cluster generation, the library preparations were sequenced on an Illumina

Hiseq 2500/2000 platform and 50bp.

Data analysis

Quality control

Raw data (raw reads) of fastq format were firstly processed through custom perl and python scripts. In this step, clean datas(clean reads) were obtained by removing reads containing ploy-N, with 5' adapter contaminants, without 3' adapter or the insert tag, containing ploy A or T or G or C and low quality reads from raw data. At the same time, Q20, Q30, and GC-content of the raw datas were calculated. Then, chose a certain range of length from clean reads to do all the downstream analyses.

Reads mapping to the reference sequence

The small RNA tags were mapped to reference sequence by Bowtie (Langmead et al., 2009) without mismatch to analyze their expression and distribution on the reference.

Known miRNA alignment

Mapped small RNA tags were used to looking for known miRNA. miRBase20.0 was used as reference, modified software mirdeep2(Friedlander et al., 2011) and srna-tools-cli were used to obtain the potential miRNA and draw the secondary structures. Custom scripts were used to obtain the miRNA counts as well as base bias on the first position of identified miRNA with certain length and on each position of all identified miRNA respectively.

Remove tags from these sources

To remove tags originating from protein-coding genes, repeat sequences, rRNA, tRNA, snRNA, and snoRNA, small RNA tags were mapped to RepeatMasker, Rfam database or those types of datas from the specified species itself.

Novel miRNA prediction

The characteristics of hairpin structure of miRNA precursor can be used to predict novel miRNA. The available software miREvo (Wen et al., 2012) and mirdeep2 (Friedlander et al., 2011) were integrated to predict novel miRNA through exploring the secondary structure, the Dicer cleavage site and the minimum free energy of the small RNA tags unannotated in the former steps. At the same time, custom scripts were used to obtain the identified miRNA counts as well as base bias on the first position with certain length and on each position of all identified miRNA respectively.

Small RNA annotation summary

Summarizing all alignments and annotations obtained before. In the alignment and annotation before, some small RNA tags may be mapped to more than one category. To make every unique small RNA mapped to only one annotation, we follow the following priority rule: known miRNA > rRNA > tRNA > snRNA > snoRNA > repeat > gene > NAT-siRNA > gene > novel miRNA > ta-siRNA. The total rRNA proportion was used a marker as sample quality indicator. Usually it should be less than 60% in plant samples and 40% in animal samples as

high quality.

miRNA editing analysis

Position 2~8 of a mature miRNA were called seed region which were highly conserved. The target of a miRNA might be different with the changing of nucleotides in this region. In our analysis pipeline, miRNA which might have base edit could be detected by aligning all the sRNA tags to mature miRNA, allowing one mismatch.

miRNA family analysis

Exploring the occurrence of miRNA families identified from the samples in other species. In our analysis pipeline, known miRNA used miFam.dat (http://www.mirbase.org/ftp.shtml) to look for families; novel miRNA precursor was submitted to Rfam (http://rfam.sanger.ac.uk/search/) to look for Rfam families.

Target gene prediction

Predicting the target gene of miRNA was performed by miRanda (Enright et al, 2003) for animals.

Quantification of miRNA

miRNA expression levels were estimated by TPM (transcript per million) through the following criteria (Zhou et al., 2010): Normalization formula: Normalized expression = mapped readcount/Total reads*1000000

Differential expression of miRNA

For the samples with biological replicates: Differential expression analysis of two conditions/groups was performed using the DESeq R package (1.8.3). The P-values was adjusted using the Benjamini& Hochberg method. Corrected P-value of 0.05 was set as the threshold for significantly differential expression by default.

mRNA Methods

RNA isolation, quantification and qualification

RNA degradation and contamination was monitored on 1% agarose gels. RNA purity was checked using the Nano Photometer® spectrophotometer (IMPLEN, CA, USA). RNA concentration was measured using Qubit® RNA Assay Kit in Qubit® 2.0 Flurometer (Life Technologies, CA, USA). RNA integrity was assessed using the RNA Nano 6000 Assay Kit of the Bioanalyzer 2100 system (Agilent Technologies, CA, USA).

Library preparation for lncRNA sequencing

A total amount of 3 μg RNA per sample was used as input material for the RNA sample preparations. Firstly, ribosomal RNA was removed by Epicentre Ribo-zero™ rRNA Removal Kit (Epicentre, USA), and rRNA free residue was cleaned up by ethanol precipitation. Subsequently, sequencing libraries were generated using the rRNA-depleted RNA by

NEBNext® Ultra™ Directional RNA Library Prep Kit for Illumina® (NEB, USA) following manufacturer's recommendations. Briefly, fragmentation was carried out using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer(5X). First strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase(RNaseH-). Second strand cDNA synthesis was subsequently performed using DNA Polymerase I and RNase H. In the reaction buffer, dNTPs with dTTP were replaced by dUTP. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with hairpin loop structure were ligated to prepare for hybridization. In order to select cDNA fragments of preferentially 150~200 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, USA). Then 3 μl USER Enzyme (NEB, USA) was used with size-selected, adaptor-ligated cDNA at 37° C for 15 min followed by 5 min at 95°C before PCR. Then PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers and Index (X) Primer. At last, products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system.

Clustering and sequencing

The clustering of the index-coded samples was performed on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumia) according to the manufacturer's instructions. After cluster generation, the libraries were sequenced on an Illumina Hiseq 4000 platform and 150 bp paired-end reads were generated.

Data analysis

Quality control

Raw data(raw reads) of fastq format were firstly processed through in-house perl scripts. In this step, clean data(clean reads) were obtained by removing reads containing adapter, reads on containing ploy-N and low quality reads from raw data. At the same time, Q20, Q30 and GC content of the clean data were calculated. All the down stream analyses were based on the clean data with high quality.

Mapping to the reference genome

Reference genome and gene model annotation files were downloaded from genome website directly. Index of the reference genome was built using bowtie2 v2.2.8 and paired-end clean reads were aligned to the reference genome using HISAT2(Langmead, B.et al) v2.0.4. HISAT2 was run with '--rna-strandness RF', other parameters were set as default.

Transcriptome assembly

The mapped reads of each sample were assembled by StringTie (v1.3.1) (Mihaela Pertea.et al. 2016) in a reference-based approach. StringTie uses a novel network flow algorithm as well as an optional de novo assembly step to assemble and quantitate full-length transcripts representing multiple splice variants for each gene locus.

Coding potential analysis

CNCI

CNCI (Coding-Non-Coding-Index) (v2) profiles adjoining nucleotide triplets to effectively distinguish protein-coding and non-coding sequences independent of known annotations (Sun et al. 2013). We use CNCI with default parameters.

CPC

CPC (Coding Potential Calculator) (0.9-r2) mainly through assess the extent and quality of the ORF in a transcript and search the sequences with known protein sequence database to clarify the coding and non-coding transcripts (Kong et al. 2007). We used the NCBI eukaryotes' protein database and set the e-value '1e-10' in our analysis.

Pfam-sca

We translated each transcript in all three possible frames and used Pfam Scan (v1.3) to identify occurrence of any of the known protein family domains documented in the Pfam database (release 27; used both Pfam A and Pfam B)（Punta, et al. 2012）. Any transcript with a Pfam hit would be excluded in following steps. Pfam searches use default parameters of -E 0.001 --domE 0.001 (Bateman, et al. 2002).

phyloCSF

PhyloCSF (phylogenetic codon substitution frequency) (v20121028) examines evolutionary signatures characteristic to alignments of conserved coding regions, such as the high frequencies of synonymous codon substitutions and conservative amino acid substitutions,

and the low frequencies of other missense and non-sense substitutions to distinguish protein-coding and non-coding transcripts (Lin et al. 2011). We build multi-species genome sequence alignments and run phyloCSF with default parameters. Transcripts predicted with coding potential by either/all of the four tools above were filtered out, and those without coding potential were our candidate set of lncRNAs.

Conservative analysis

Phast (v1.3) is a software package contains much of statistical programs, most used in phylogenetic analysis (Siepel, et al. 2005), and phastCons is a conservation scoring and identificating program of conserved elements. We used phyloFit to compute phylogenetic models for conserved and non-conserved regions among species and then gave the model and HMM transition parameters to phyloP to compute a set of conservation scores of lncRNA and coding genes.

Quantification of gene expression level

Cuffdiff (v2.1.1) was used to calculate FPKMs of both lncRNAs and coding genes in each sample (Trapnell, C. et al. 2010). Gene FPKMs were computed by summing the FPKMs of transcripts in each gene group. FPKM means fragments per kilo-base of exon per million fragments mapped, calculated based on the length of the fragments and reads count mapped to this fragment.

Differential expression analysis

The Ballgown suite includes functions for interactive exploration of the transcriptome assembly, visualization of transcript structures and feature-specific abundances for each locus, and post-hoc annotation of assembled features to annotated features(Alyssa C. Frazee et al.2014). Transcripts with an P-adjust <0.05 were assigned as differentially expressed. Cuffdiff provides statistical routines for determining differential expression in digital transcript or gene expression data using a model based on the negative binomial distribution (Trapnell, C. et al. 2010). Transcripts with an P-adjust <0.05 were assigned as differentially expressed.

References

Anders, S., Huber, W. (2010). Differential expression analysis for sequence count data. Genome Biology, doi:10.1186/gb-2010-11-10-r106. (DESeq)

Alyssa C. Frazee, Geo Pertea, Andrew E. Jaffe, Ben Langmead, Steven L. Salzberg

Enright AJ, John B, Gaul U, Tuschl T, Sander C, et al. (2003) MicroRNA targets in Drosophila. Genome Biol 5: R1. (miRanda)

Friedlander M.R., Mackowiak S.D., Li N., Chen W., Rajewsky N. (2011). miRDeep2 accurately identifies known and hundreds of novel microRNA genes in seven animal clades. Nucleic Acids Res 40:37-52. (miRDeep2)

Jeffrey T. Leek.(2014) Flexible analysis of transcriptome assemblies with Ballgown. Biorxiv.

Kanehisa, M., M. Araki, et al. (2008). KEGG for linking genomes to life and the Langmead, B., Trapnell, C., Pop, M., & Salzberg, S. L. (2009). Ultrafast and memory-efficient alignment

of short DNA sequences to the human genome. Genome Biol, 10(3), R25. (Bowtie)

Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, et al. (2008). KEGG for linking genomes to life and the environment. Nucleic Acids research36: D480–484. (KEGG)

Langfelder, P.,Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. (coexpression)

Mao, X., Cai, T., Olyarchuk, J.G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG orthology (KO) as a controlled vocabulary. Bioinformatics 21, 3787–3793. (KOBAS)

McKenna, A, Hanna, M, Banks, E, Sivachenko, A, Cibulskis, K, Kernytsky, A, Garimella, K, Altshuler, D, Gabriel, S, Daly, M, DePristo, MA. (2010). The Genome Analysis Toolkit: A MapReduce framework for analyzing next-generation DNA sequencing data. Genome Research. (GATK)

Siepel, A., Bejerano, G., Pedersen, J.S., et al. (2005). Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes. Genome Res. 15, 1034-1050. (Phast)

Storey, J. D. (2003). The positive false discovery rate: A Bayesian interpretation and the q-value, Annals of Statistics. 31: 2013-2035. (qvalue)

Trapnell, C. et al. (2010). Transcript assembly and quantification by RNA-seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat. Biotechnol. (Cufflinks)

Wang L., Feng Z., Wang X., Wang X., Zhang X. (2010). DEGseq: An R package for identifying differentially expressed genes from RNA-seq data. Bioinformatics 26, 136-8. (DEGseq)

Wen M., Shen Y., Shi S., and Tang T. (2010). miREvo: An Integrative microRNA Evolutionary Analysis Platform for Next-generation Sequencing Experiments. BMC Bioinformatics 13:140. (miREvo)

Wu HJ, Ma YK, Chen T, Wang M, Wang XJ (2012) PsRobot: a web-based plant small RNA meta-analysis toolbox. Nucleic Acids Res 40: W22–W28.(psRobot)

Young, M.D., Wakefield, M.J., Smyth, G.K., and Oshlack, A (2010). goseq: Gene Ontology testing for RNA-seq datasets. (goseq)

Zhou L., Chen J., Li Z., Li X., Hu X., et al. (2010). Integrated profiling of microRNAs and mRNAs: microRNAs located on Xq27.3 associate with clear cell renal cell carcinoma. PLoS One 5: e15224. (TPM)