



Transcriptomic Analyses for Identification and Prioritization of Genes Associated With Alzheimer's Disease in Humans

Yuchen Shi^{††}, Hui Liu^{††}, Changbo Yang^{††}, Kang Xu^{††}, Yangyang Cai¹, Zhao Wang¹, Zheng Zhao^{2*}, Tingting Shao^{1*} and Yixue Li^{1*}

¹ College of Bioinformatics Science and Technology, Harbin Medical University, Harbin, China, ² Beijing Neurosurgical Institute, Capital Medical University, Beijing, China

OPEN ACCESS

Edited by:

Meng Zhou,
Wenzhou Medical University, China

Reviewed by:

Yucheng Yang,
Yale University, United States
Qi Zhao,
Shenyang Aerospace University,
China

*Correspondence:

Yixue Li
biolyx@163.com
Tingting Shao
shaotingting_05@163.com
Zheng Zhao
zhaozheng0503@sina.com

^{††}These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Bioinformatics and Computational
Biology,
a section of the journal
Frontiers in Bioengineering and
Biotechnology

Received: 11 October 2019

Accepted: 14 January 2020

Published: 21 February 2020

Citation:

Shi Y, Liu H, Yang C, Xu K, Cai Y,
Wang Z, Zhao Z, Shao T and Li Y
(2020) Transcriptomic Analyses for
Identification and Prioritization of
Genes Associated With Alzheimer's
Disease in Humans.
Front. Bioeng. Biotechnol. 8:31.
doi: 10.3389/fbioe.2020.00031

Long non-coding RNAs (lncRNAs), as important ncRNA regulators, play crucial roles in the regulation of various biological processes, and their aberrant expression is related to the occurrence and development of diseases, which is gradually validated by more and more studies. Alzheimer's disease (AD) is a chronic neurodegenerative disease that often develops slowly and gradually deteriorates over time. However, which functions the lncRNAs perform in AD are almost unknown. In this study, we performed transcriptome analysis in AD, containing 12,892 known lncRNAs and 19,053 protein-coding genes (PCGs). Further, 14 down-regulated and 39 up-regulated lncRNAs were identified, compared with normal brain samples, which indicated that these lncRNAs might play critical roles in the pathogenesis of AD. In addition, 19 down-regulated and 28 up-regulated PCGs were also detected. Using the differentially expressed lncRNAs and PCGs through the WGCNA method, an lncRNA–mRNA co-expressed network was constructed. The results showed that lncRNAs RP3-522J7, MIR3180-2, and MIR3180-3 were frequently co-expressed with known AD risk PCGs. Interestingly, PCGs in the network are significantly enriched in brain- or AD-related biological functions, including the brain renin–angiotensin system, cell adhesion, neuroprotective role of THOP1 in AD, and so on. Furthermore, it was shown that 18 lncRNAs and 7 PCGs were highly expressed in normal brain tissue relative to other normal tissue types, suggesting their potential as diagnostic markers of AD, especially RP3-522J7, MIR3180-2, MIR3180-3, and CTA-929C8. In total, our study identified a compendium of AD-related dysregulated lncRNAs and characterized the corresponding biological functions of these lncRNAs in AD, which will be helpful to understand the molecular basis and pathogenesis of AD.

Keywords: long non-coding RNA, Alzheimer's disease, transcriptomic analyses, RNA-seq, differential expression analysis

INTRODUCTION

Alzheimer's disease (AD) is an insidious progressive neurodegenerative disease, which is affected by many factors, including environmental factors and genetic and epigenetic variations (Toledo et al., 2015), but there is no doubt that aging is the biggest risk factor (Kold-Christensen and Johannsen, 2019). As the global population ages, there will be two billion people aged 60 years or over in the

world by 2050 (Harper, 2014), and two new patients will be diagnosed with AD every minute, resulting in one million new patients who will be diagnosed every year (Alzheimer's Association, 2015). Although the pathophysiology of AD still remains unclear (Xu et al., 2018), and even if there are no effective treatment methods currently, it is generally accepted that only the earliest intervention is likely to affect the progression of the disease (Sood et al., 2015).

With the development of life science research, it has been found that the human transcriptome is more complicated than previously known. Indeed, a considerable part of the human genome regions could transcribe as long non-coding RNAs (lncRNAs), a set of non-coding transcripts longer than 200 nt, which play important roles in various life activities, such as embryonic development, cell differentiation, aging, and complex diseases (Derrien et al., 2012; Rinn and Chang, 2012). Many studies have proved that lncRNAs are involved in many physiological and pathological processes through transcriptional or post-transcriptional regulatory mechanisms and thus play an important role in the process of the whole life, so it has become a hot spot of genetic and epigenetic research (Mercer and Mattick, 2013; Papait et al., 2013; Shi et al., 2013; Fatica and Bozzoni, 2014; Sultmann and Diederichs, 2014). As previous research has shown, due to the complex and transcriptional regulation mechanisms, lncRNAs could be focused on their roles in human complex diseases, including AD and cancers, to gain new insights into complex disease pathways, to identify biomarkers to improve diagnostic accuracy, and to examine the impact of treatment (Zhou et al., 2019). High-throughput sequencing technology has significant advantages, providing opportunities for insight into the genomic and transcriptomic research with large data sets, and it could be helpful to dissect the comprehensive transcriptome characterization of complex diseases (Sultan et al., 2008; Xiong et al., 2012; Prasad et al., 2016). However, the role of lncRNAs in the pathogenesis and progression of certain human complex diseases remains unclear, particularly AD. By comparing a ck-p25 AD model and control samples in mice, it was found that histone modification could regulate the differential expression of lncRNAs (Wan et al., 2016). Thus, it is indicated that lncRNAs might be helpful in investigating the transcription landscape of brain tissue in AD.

In this study, we aimed to identify the gene expression patterns in AD patients and controls, identify the lncRNAs involved in the AD-related dysregulated biological processes, and then to prioritize important lncRNAs based on the co-expression networks composed of differentially expressed lncRNAs and protein-coding genes (PCGs) in AD. Our results suggest that dysregulated expressed transcripts could affect complex disease, including lncRNAs and PCGs. Differentially expressed lncRNAs are significantly co-expressed with PCGs, which could be constructed in the co-expressed network to exhibit the complex regulatory relations in AD. Functional analysis revealed that lncRNAs might be involved in the AD-related biological processes, even regulate the AD-related function modules. Summarizing the above, our study provided a new set of AD-related dysregulated lncRNAs and identified the corresponding

biological functions of these lncRNAs in AD. It will be helpful for further understanding the pathogenesis for AD.

MATERIALS AND METHODS

RNA-seq Data Analysis

NGS data used in this research were derived from public resource GEO, including RNA-seq data of human cerebral cortex tissues from nine patients and eight controls (GEO accession GSE53697) (Scheckel et al., 2016). The RNA-seq data were mapped into the human reference genome (Hg38 version) by Tophat2 (Kim et al., 2013). For the measurement of expression values for lncRNAs and PCGs, we used Cufflinks v2.1.1 with FPKM (Trapnell et al., 2012).

Differential Expression Analysis

Wilcoxon rank sum test was used to identify the differentially expressed lncRNAs and PCGs between the AD cases and the control group with the threshold of $P < 0.01$ (Han et al., 2019). Classification of lncRNAs and PCGs was according to the gene annotation file derived from GENCODE. The following biotypes are considered as known lncRNAs: "3prime_overlapping_ncrna," "ambiguous_orf," "antisense," and "antisense_RNA," "lincRNA," "ncrna_host," "non_coding," "non_stop_decay," "processed_transcript," "retained_intron," "sense_intronic," and "sense_overlapping." And the biotype of "protein_coding" is considered as a PCG.

Collection of Aging-Related Gene Data Set

To investigate the relationship between differentially expressed genes in AD and aging, a list of aging-related genes was collected, which were defined as the genes involved in the development or aging-associated GO biological processes. And the functional annotation for genes and GO terms was derived from the Gene2GO file in NCBI¹. And then, a cumulative hypergeometric test was used to investigate whether the AD-associated differentially expressed PCGs were enriched in the aging gene set with statistical significance.

Co-expression Network Analysis

As important regulators, lncRNAs are considered to be involved the corresponding biological processes by regulating their target genes. For each differentially expressed lncRNA, we attempted to identify its regulatory target genes. The co-expression network for differentially expressed lncRNAs and PCGs was constructed by the WGCNA method with empirical threshold with the value equal to nine (Langfelder and Horvath, 2008).

Gene Set Enrichment Analysis

To identify biological processes and cellular components which would be regulated by aberrantly expressed lncRNAs and PCGs in AD, enrichment analyses were performed for each lncRNA target

¹<https://www.ncbi.nlm.nih.gov/>

PCG using the R package TCGAbiolinks (Colaprico et al., 2016), and GO/KEGG terms with adjusted p value < 0.01 by Benjamini-Hochberg methods were considered.

Brain-Elevated Expression Analysis

The tissue-specific RNA-seq data set was obtained from the Genotype-Tissue Expression (GTEx, 2013) project, in which gene-level average RPKM values are reported for each tissue sample across 30 tissues including the brain (2013). A certain differentially expressed lncRNA or PCG would be defined as brain elevated if this transcription's expression value in brain tissue is more than five times compared with the average value in all other tissues.

AD-Related miRNA

To prioritize AD-related miRNA, we used the miRanda algorithm and searched the starBase database to obtain the mRNA-miRNA targeting relationship (Betel et al., 2010; Li et al., 2014). AD-related miRNAs were collected from the HMDD database, which manually curated the experimentally validated human disease-associated miRNA information (Huang et al., 2019). We analyzed the target relationship between AD-related miRNAs and differentially expressed PCGs and then ranked candidate miRNAs according to the count of the overlap between the differentially expressed PCGs targeted by known AD miRNAs and the targets of each miRNA.

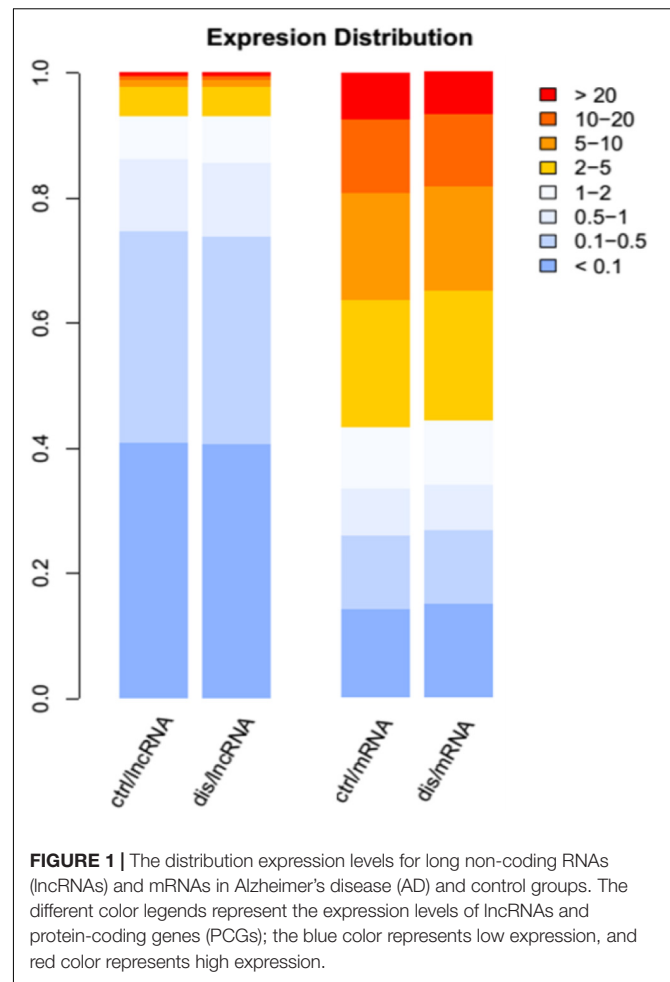
RESULTS

Transcriptome Analysis in AD

A total of 17 RNA libraries were prepared from human brain samples, including nine AD and eight controls. From these samples, a total of 1,014,513,141 read pairs were generated from RNA-seq experiments, and >700 million read pairs (72.4%) were aligned to the human genome (Hg38 version). For the measurement of expression levels for lncRNAs and PCGs, we used Cufflinks. The gene expression levels were compared between AD and control samples, including lncRNAs and PCGs. The result revealed that about 12,892 lncRNAs and 19,053 PCGs were expressed. Consistently with previous studies, the expression levels of PCGs were generally higher than those of lncRNAs. In detail, there were just about 60% of lncRNAs with the FPKM value above 0.1, whereas the majority of PCGs were expressed above one (Figure 1). Moreover, 54.5% of lncRNAs were expressed in more than 15 samples and 1.6 times enriched for PCGs compared with lncRNAs (Supplementary Figure S1). These results revealed that lncRNAs were expressed widely in AD, indicating their potential important roles in the development of disease.

Differential Expression of Both lncRNAs and PCGs in AD

Previous studies have confirmed that the variant events of the human brain transcriptome are involved in the pathophysiological process of AD. Thus, the identification of



the aberrantly expressed genes in AD patients might be helpful to understand the molecular mechanism of AD (Twine et al., 2011; Bernstein et al., 2016; Hensman Moss et al., 2017). We re-analyzed AD transcriptome data derived from a previous independent study which focused on ELAV-like protein binding to genes in the human brain yet did not analyze differential expression analysis of lncRNAs and PCGs (Scheckel et al., 2016). In this part, we removed the lncRNAs overlapping with PCGs in order to ensure the accuracy of the results. We found 100 differentially expressed genes, including 39 up-regulated lncRNAs and 14 down-regulated lncRNAs, as well as 28 up-regulated PCGs and 19 down-regulated PCGs (p -value < 0.01) (Supplementary Table S1). The unsupervised hierarchical clustering results revealed that differentially expressed lncRNAs (Figure 2A) and PCGs (Supplementary Figure S2) could well distinguish AD samples from control samples, indicating that there might be existing distinct signatures in expression level for AD. Notably, several lncRNAs, including CTA-929C8, RP11-461L13, and PSMG3-AS1, have relatively high expression levels in AD or control samples. For example, the expression value of CTA-929C8 in normal tissue is 0.3665 FPKM but decreases to 0.069 FPKM in AD. Another lncRNA is PSMG3-AS1,

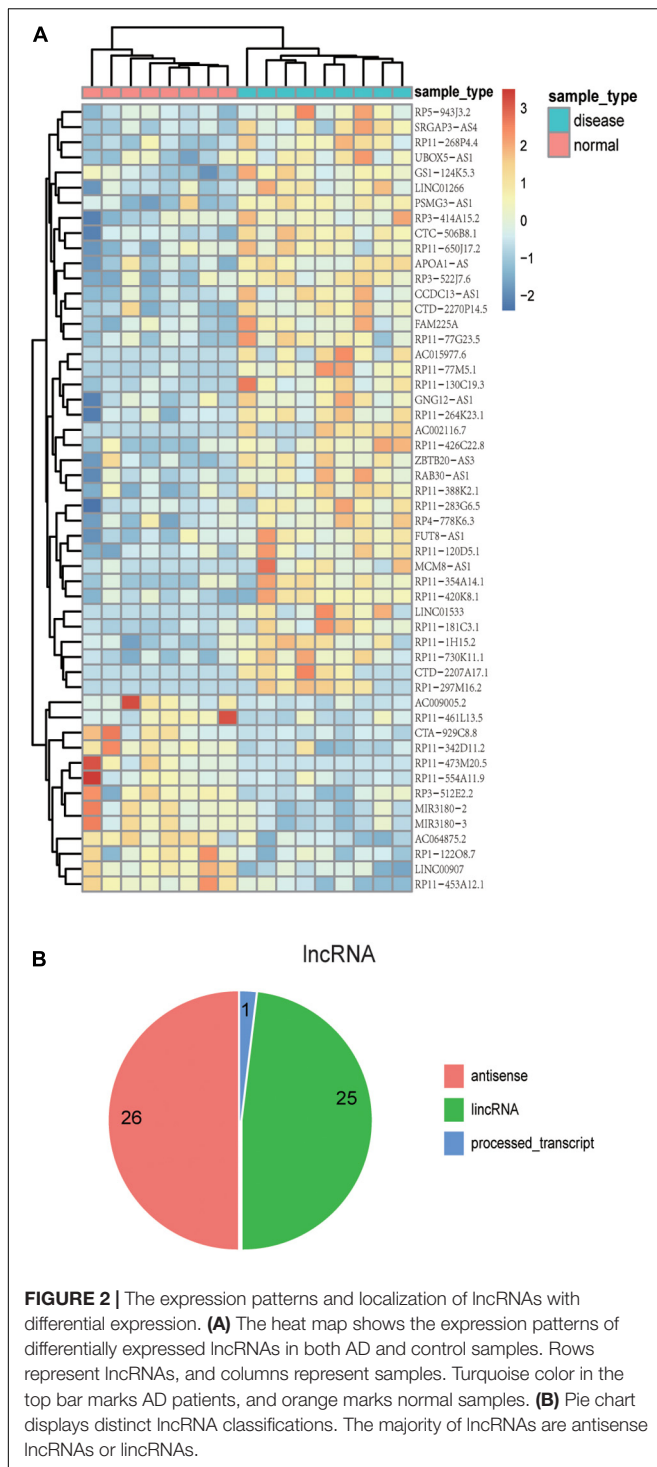


FIGURE 2 | The expression patterns and localization of lncRNAs with differential expression. **(A)** The heat map shows the expression patterns of differentially expressed lncRNAs in both AD and control samples. Rows represent lncRNAs, and columns represent samples. Turquoise color in the top bar marks AD patients, and orange marks normal samples. **(B)** Pie chart displays distinct lncRNA classifications. The majority of lncRNAs are antisense lncRNAs or lincRNAs.

which is up-regulated with the mean expression value of 1.718 FPKM in AD.

Further, among the differentially expressed PCGs in AD, eight genes have been validated to be associated with AD, including VTI1A, CUX1, S100B, AGT, CD44, NTS, IRAK4, and AQP1, which are reported by the DisGeNET database (Pinero et al., 2017). The first two genes are down-regulated,

and the others are up-regulated in AD. Moreover, VTI1A was validated to have susceptibility loci for late-onset AD (Grube et al., 2006), even in glioma, based on a genome-wide association study (Kinnersley et al., 2015). As another example, CD44, a surface antigen expressed across multiple tissues, was up-regulated in AD samples, which corresponds with previous studies (Uberti et al., 2010). The results indicated that CD44 could play crucial roles in driving the immune response into infected tissues in the central nervous system. The aberrantly expressed S100B could seriously affect one of the hallmarks of AD and neuroinflammation; thus, it plays an important role in the pathophysiological process of AD, and the drugs targeting S100B might have a great impact on the treatment of this disease (Cirillo et al., 2015). These results indicate that the differentially expressed lncRNAs and PCGs are significantly associated with AD, which might affect the occurrence and development of AD.

Moreover, we investigated the genomic localization of differentially expressed lncRNAs compared with PCGs. As the results show in **Figure 2B**, 26 differentially expressed lncRNAs are antisense lncRNAs, and 25 are lincRNAs. Next, we computed the distance between differentially expressed lncRNAs and differentially expressed PCGs. As a result, we found that there are several lncRNAs located near important PCGs in the human genome. For example, six differentially expressed lncRNAs are located within 10 MB of S100B, and five of them are up-regulated in AD corresponding with S100B. The intergenic lncRNA, RP3-522J7, is an up-regulated lncRNA in AD with a fold change value of 2.65, compared with control samples, whose genomic distance is 3.3 MB apart from S100B. Therefore, it is concluded that the lncRNA PR3-522J7 might cis-regulate the expression of S100B. Another example is the gene CD44, which had 11 lncRNAs nearby, with a distance less than 10 Mb. Especially among the 11 neighbor lncRNAs, two lncRNAs described above, CTA-929C8 and CCDC13-AS1, are 7.8 and 7.5 MB apart from CD44. Therefore, we concluded that these two lncRNAs might cis-regulate the expression of CD44. These results indicate that some of these differentially expressed lncRNAs might cis-regulate the known AD protein-coding genes to influence the occurrence and development of AD.

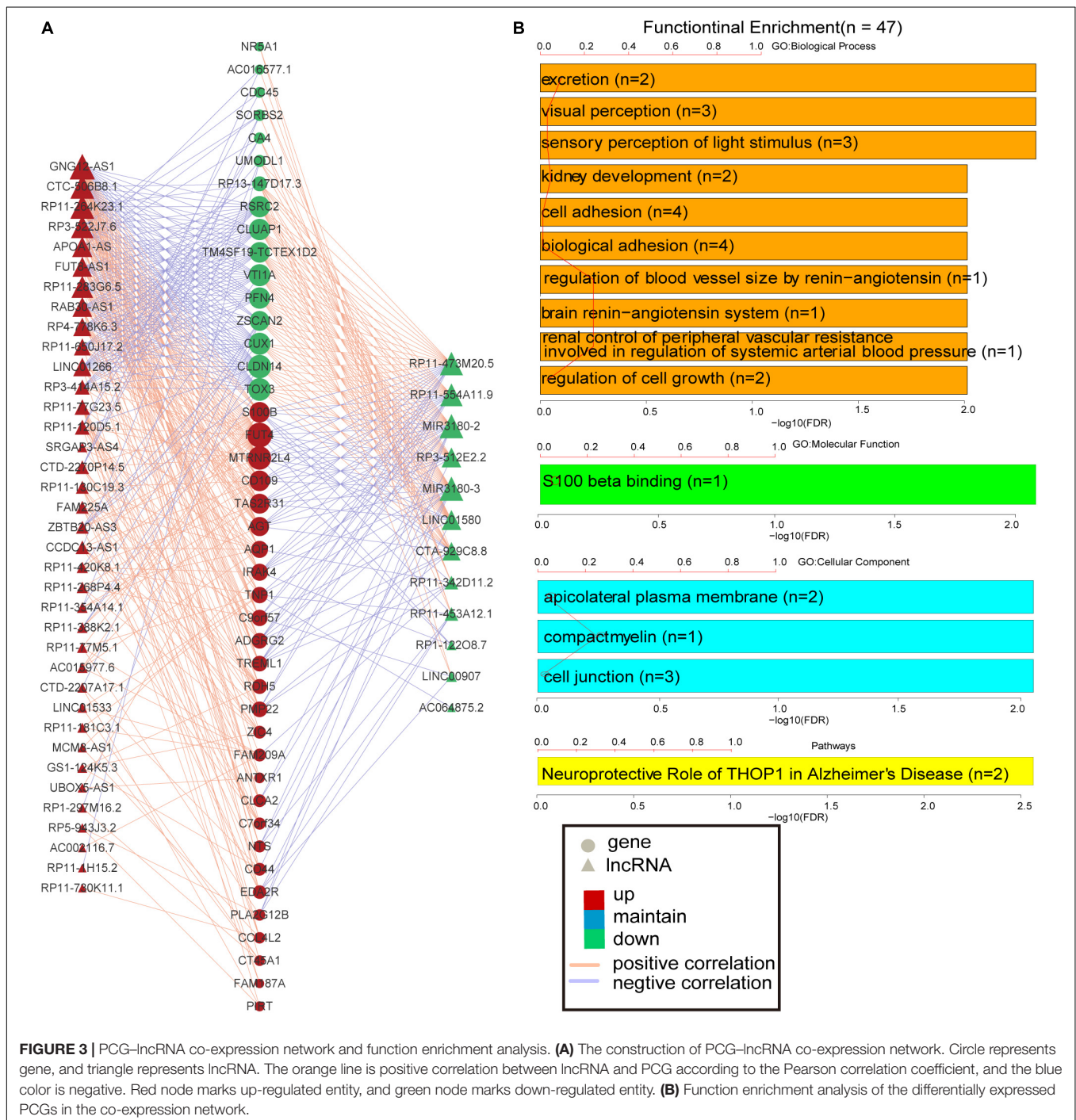
Additionally, 92 development or age-related GO terms were collected to obtain a total of 2,435 age-related genes (**Supplementary Table S2**), and then, we found that 11 of these differentially expressed genes were identified as aging-related genes (**Supplementary Table S3**). The results of a cumulative hypergeometric test indicated that differentially expressed genes in diseases were significantly correlated with aging with statistical significance ($p = 0.004650242$), including AGT and CUX1, which are involved in function “aging” and “multicellular organism development,” respectively. Moreover, some differentially expressed genes are labeled as aging or longevity by the HAGR database and some studies. For example, S100B and AGT were labeled as aging genes in the HAGR database, while TOX3 was considered as a longevity gene (Tacutu et al., 2018). Another example is CLUAP1, which has been shown to vary significantly with age (Peters et al., 2015). These results indicate that these differentially expressed PCGs are associated

with aging, confirm the effect of aging on AD, and are consistent with clinical observations.

Co-expression Network Analysis Reveals Differentially Expressed Genes Involved in AD-Related Biological Functions

More and more lncRNAs have been validated to play important roles in disease-associated biological processes (Sun et al., 2014;

Zhou et al., 2017); however, there are still a considerable number of lncRNAs uncovered involved in the physiological process during embryonic development and cell differentiation, even in the pathological processes of human complex diseases. Based on the knowledge of the biological functions of the coding gene, the biological processes of lncRNAs could be predicted, with the hypothesis that if the expression of a certain lncRNA is correlated with the lists of PCGs, this gene co-expression pattern may provide reliable evidence that they are involved



in the same or similar biological functions. The lncRNA-PCG co-expression network for differentially expressed lncRNAs and PCGs was constructed via R package WGCNA. As a result, the co-expression network was composed of 489 links between 49 differentially expressed lncRNAs and 44 PCGs (Figure 3A). These lncRNAs and PCGs are tightly connected together. Especially, S100B is co-expressed with 18 lncRNAs, including seven down-regulated lncRNAs and 11 up-regulated lncRNAs. For example, the weight value of the link between lncRNA RP3-522J7 and

PCG S100B is 0.043867. Moreover, this lncRNA is located within 3.3 Mb from S100B. Therefore, we concluded that RP3-522J7 might cis-regulate the expression of S100B. In addition, we found that MIR3180-2 and MIR3180-3 shared relatively more co-expressed PCGs, including VT11A, CUX1, S100B, AGT, NTS, and IRAK4, which are known as AD-related PCGs. Both the two lncRNAs are the host genes of hsa-miR-3180-3p and hsa-miR-3180-5p; thus, it was proposed that these two miRNAs might be up-regulated in AD.

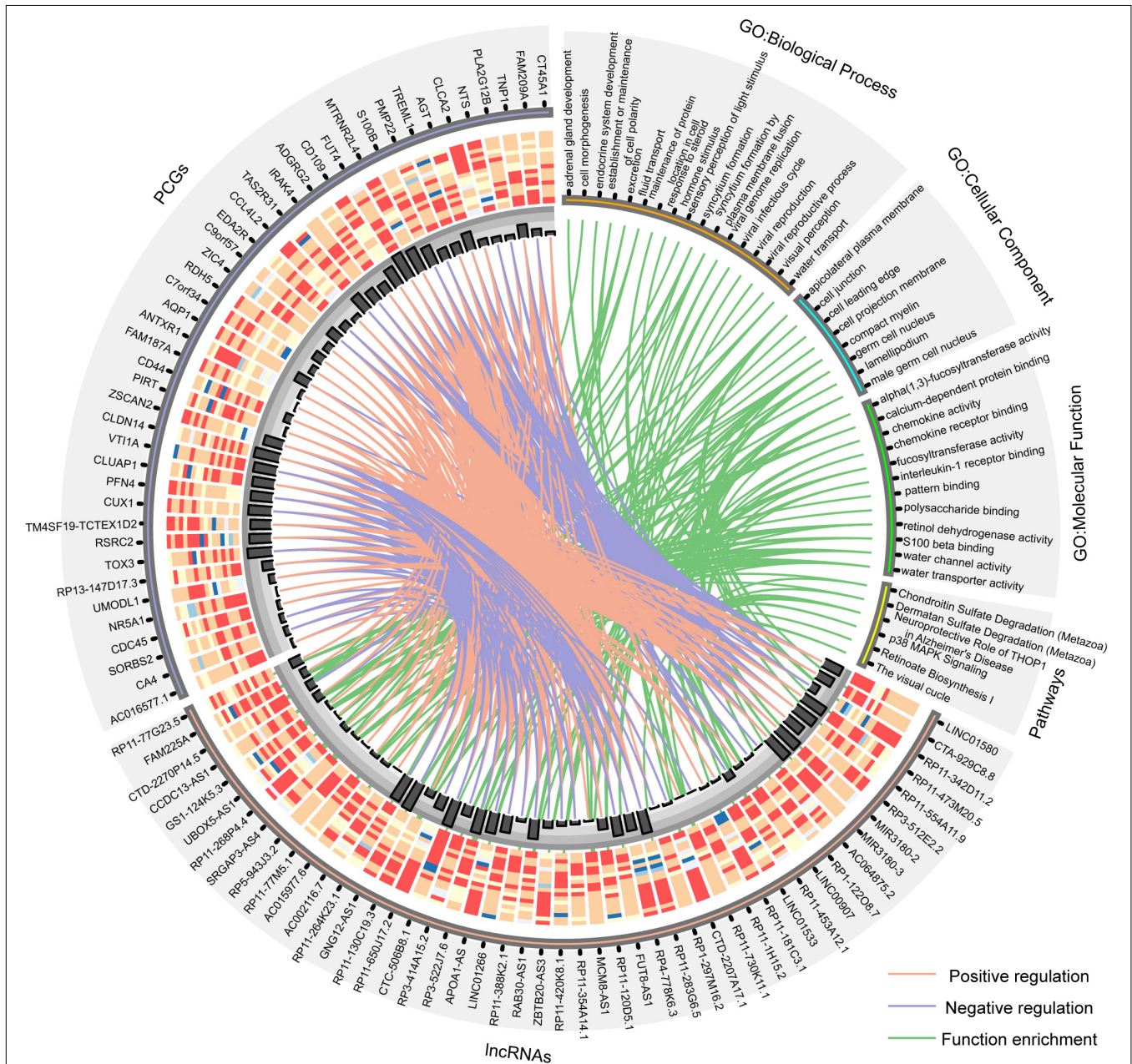


FIGURE 4 | Circos plot shows the relationship between lncRNAs, targeted PCGs, and parts of functions (FDR < 0.005). The orange line is positive correlation between lncRNA and PCG according to the Pearson correlation coefficient, the blue color is negative, and the green lines connect lncRNAs and the functions that their targeted PCGs are enriched in. Bar plot shows the number of target relationships.

Furthermore, the GO and KEGG function enrichment analysis was performed for the differentially expressed PCGs in the co-expression network by the TCGAbiolinks tool (Xiong et al., 2012). **Figure 3B** and **Supplementary Table S4** show the over-represented terms and pathways, among the most enriched of which are those involved in cell adhesion (FDR = 0.00969), regulation of cell growth (FDR = 0.00969), visual perception (FDR = 0.00460), brain renin-angiotensin system (FDR = 0.00969), neuroprotective role of THOP1 in AD (FDR = 0.00275), and so on. It's worth noting that many PCGs targeted by differentially expressed lncRNAs in AD are involved in AD- and aging-related biological functions (**Supplementary Table S5**), such as CTA-929C8 target PCGs involved in the brain renin-angiotensin system (FDR = 0.00969), regulation of inflammatory response (FDR = 0.00969), and regulation of vasoconstriction (FDR = 0.00969) (**Supplementary Figure S3**); RP3-522J7.6 target PCGs involved in the brain renin-angiotensin system (FDR = 0.00944), the renal system process involved in the regulation of systemic arterial blood pressure (FDR = 0.00944), and the neurological system process involved in regulation of systemic arterial blood pressure (FDR = 0.00944) (**Supplementary Figure S4**); and MIR3180-2 and MIR3180-3 target PCGs involved in the neuroprotective role of THOP1 in AD (FDR = 0.00130) and blood vessel remodeling (FDR = 0.00944) (**Supplementary Figures S5, S6**). The relationship between lncRNAs, targeted PCGs, and parts of functions are shown in **Figure 4** (FDR < 0.005). All these categories are associated with brain-related functions and are implicated in AD or aging. Therefore, these differentially expressed lncRNAs might be involved in AD and aging via regulating these molecular functions.

Prioritization of Candidate lncRNAs as Potential Diagnostic Markers of AD

Further, we investigated whether these differentially expressed lncRNAs and PCGs tend to be particularly highly expressed in human cerebral tissues, which might be involved in brain-related biological processes and could be diagnostic markers of AD. We obtained high-throughput data from the GTEx project including 30 human tissues, and then the mean expression of the differentially expressed lncRNAs and PCGs was calculated (2013) (**Figure 5A** and **Supplementary Figure S3**). Eighteen differentially expressed lncRNAs are specifically highly expressed in cerebral tissue, including the lncRNAs described above, such as RP3-522J7, CCDC13-AS1, MIR3180-2, MIR3180-3, CTA-929C8, and so on (**Figure 5A**). On the other hand, there are seven PCGs, including three known AD-related genes, which are S100B, AGT, and NTS (**Supplementary Figure S7**). lncRNA CTA-929C8 is highly expressed in brain tissue, and its expression level was enriched in normal brain tissue, which was more than about 1,000-fold compared with other normal tissues (**Figure 5B**). We concluded that this lncRNA might be a suitable marker in AD. In addition, for the two PCGs, S100B and PIRT, we found that they are also enriched in normal brain tissues and that their

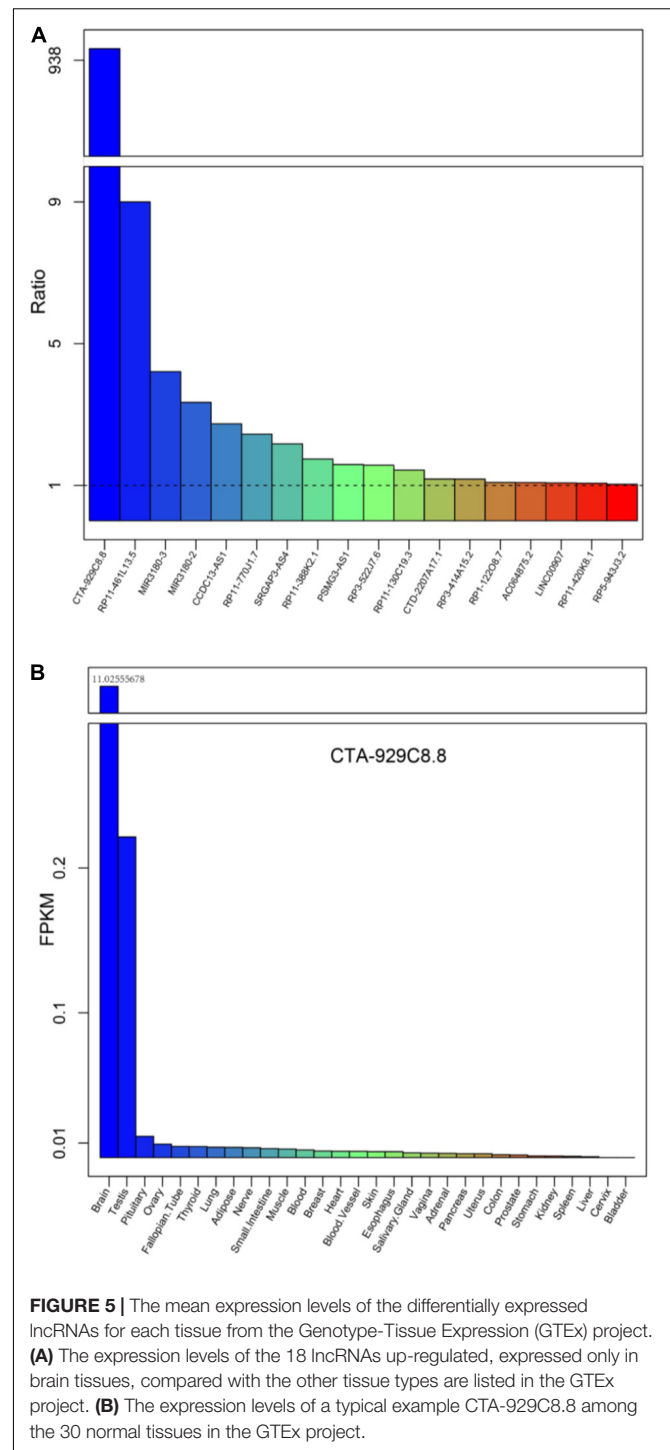


FIGURE 5 | The mean expression levels of the differentially expressed lncRNAs for each tissue from the Genotype-Tissue Expression (GTEx) project. **(A)** The expression levels of the 18 lncRNAs up-regulated, expressed only in brain tissues, compared with the other tissue types are listed in the GTEx project. **(B)** The expression levels of a typical example CTA-929C8.8 among the 30 normal tissues in the GTEx project.

expression levels were more than 10 times versus the other tissues (**Supplementary Figure S7**). Moreover, the two genes are up-regulated in AD. Therefore, it is indicated that the brain-specific highly expressed lncRNAs and PCGs might be involved in the brain-related biological processes, and even affect the brain-associated functions, and could be potential diagnostic markers for AD.

DISCUSSION

RNA-seq data of AD cortex tissue was obtained to identify the differentially expressed lncRNAs and PCGs; we established a co-expression network which could be helpful to identify a list of candidate AD-related lncRNAs. Further, these differentially expressed lncRNAs are cerebral tissue-specific, highly expressed in AD patients, and also co-expressed with AD-related PCGs. The lncRNA-PCG co-expressed network was constructed by WGCNA and was enriched in the AD-associated biological processes, such as cell adhesion, brain renin-angiotensin system, neuroprotective role of THOP1 in AD, and so on. Finally, several lncRNAs were identified, including MIR3180-2, MIR3180-3, and RP3-522J7, which were most highly co-expressed with the known AD-related genes in the network, such as S100B, AGT, NTS, and so on. These results suggested that the lncRNAs might directly regulate AD-related biological functions by targeting AD risk genes. Moreover, we further prioritized a set of differentially expressed lncRNAs and PCGs, which could be acting as potential AD markers, especially RP3-522J7, MIR3180-2, MIR3180-3, CTA-929C8, S100B, and PIRT. In summary, our study identified a compendium of lncRNAs and PCGs which have never been uncharacterized for AD and normal cerebral cortex tissue in previous research and could be prioritized as diagnostic markers for AD, which are helpful in identifying the high-risk population of AD, so as to intervene in prevention and treatment as early as possible.

Furthermore, the differentially expressed lncRNAs and PCGs identified in our study were validated by an independent data set (GEO accession GSE5281) (Liang et al., 2007). We found that 23 genes are uniformly differentially expressed. Especially, several key PCGs focused on by us are significantly differentially expressed in the test set, such as VTI1A, CUX1, S100B, AGT, CD44, and AQP1, and lncRNAs PSMG3-AS1 and FUT8-AS1. Webserver AlzData² was also used to analyze AD-related PCG (Zhang et al., 2019). We found that the differentially expressed genes are in conformity in different brain regions. And the result showed that S100B was differently expressed in the entorhinal cortex. Two PCGs, AGT and IRAK4, were expressed aberrantly in temporal cortex tissue. In addition, CD44 was differentially expressed in the entorhinal cortex, hippocampus, and temporal cortex, while another PCG AQP1 was differentially expressed in both the hippocampus and temporal cortex. These results suggested that AD-related genes may have a spatial-temporal expression pattern.

It is shown that the knowledge of the functions for the regulatory factors could be helpful for further understanding the occurrence and development of complex diseases, and the dysregulated miRNAs were validated to be associated with AD (Sadlon et al., 2019). Therefore, in our study, the AD-related miRNAs were further prioritized. By identifying the miRNA-mRNA targeting relationship using the miRanda algorithm, it was shown that 18 miRNAs regulating the differential expression of PCGs were labeled as AD-related miRNAs by HMDD, and the targeted PCGs contained seven AD-related PCGs, including

VTI1A, CUX1, S100B, AGT, CD44, IRAK4, and AQP1. Then, based on the characteristics of known AD miRNAs, we ranked the miRNAs by the overlap between the differentially expressed PCGs targeted by known AD miRNAs and the targets of each miRNA to provide clues for predicting AD miRNA candidates (**Supplementary Table S6**). The higher the rank of a certain miRNA is, the more it is possible for it to be considered as an AD-related miRNA. For example, as the second-ranked miRNA, Hsa-miR-1229 directly regulated the expression level of AD-related gene SORL1, and other targeted genes, which could be involved in the biological processes of nervous system development and neurological disease (Ghanbari et al., 2016). Moreover, this miRNA has been confirmed to dysregulate expression across the different brain tissue regions in AD patients (Puthiyedth et al., 2016). Hsa-mir-328 has been reported to be associated with a variety of neurologic disorders, such as autism spectrum disorder, Huntington's, Parkinson's, and Alzheimer's, and the biological functions of its targeting genes APP and BACE1 were validated experimentally in mouse brain tissues (Boissonneault et al., 2009; Provost, 2010; Nt et al., 2018). High-throughput experimental data from starBase were used to verify the results, and similar conclusions were obtained (see **Supplementary Text S1** and **Supplementary Table S7**). These results may be helpful for exploring the role of miRNAs in AD and providing novel insight for the study of the pathophysiology of AD.

DATA AVAILABILITY STATEMENT

The RNA-seq data used in this study were obtained from GSE53697 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE53697>), GSE5281 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE5281>), and GTEx (<https://commonfund.nih.gov/GTEx/>).

AUTHOR CONTRIBUTIONS

YL, TS, and ZZ conceived and designed the experiments. YS, HL, and CY performed the experiments and analyzed the data. YS, KX, YC, and ZW wrote the manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported by the National Natural Science Foundation of China (31601065 and 61803130). The funders had no roles in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbioe.2020.00031/full#supplementary-material>

²www.alzdata.org

REFERENCES

- Alzheimer's Association (2015). 2015 Alzheimers disease facts and figures. *Alzheimers Dement* 11, 332–384. doi: 10.1016/j.jalz.2015.02.003
- Bernstein, A. I., Lin, Y., Street, R. C., Lin, L., Dai, Q., Yu, L., et al. (2016). 5-Hydroxymethylation-associated epigenetic modifiers of Alzheimers disease modulate Tau-induced neurotoxicity. *Hum. Mol. Genet.* 25, 2437–2450. doi: 10.1093/hmg/ddw109
- Betel, D., Koppal, A., Agius, P., Sander, C., and Leslie, C. (2010). Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.* 11:R90. doi: 10.1186/gb-2010-11-8-r90
- Boissonneault, V., Plante, I., Rivest, S., and Provost, P. (2009). MicroRNA-298 and microRNA-328 regulate expression of mouse beta-amyloid precursor protein-converting enzyme 1. *J. Biol. Chem.* 284, 1971–1981. doi: 10.1074/jbc.M807530200
- Cirillo, C., Capoccia, E., Iuvone, T., Cuomo, R., Sarnelli, G., Steardo, L., et al. (2015). S100B inhibitor pentamidine attenuates reactive gliosis and reduces neuronal loss in a mouse model of Alzheimer's Disease. *Biomed. Res. Int.* 2015:508342. doi: 10.1155/2015/508342
- Colaprico, A., Silva, T. C., Olsen, C., Garofano, L., Cava, C., Garolini, D., et al. (2016). TCGAAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* 44, e71. doi: 10.1093/nar/gkv1507
- Derrien, T., Johnson, R., Bussotti, G., Tanzer, A., Djebali, S., Tilgner, H., et al. (2012). The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res.* 22, 1775–1789. doi: 10.1101/gr.132159.111
- Fatica, A., and Bozzoni, I. (2014). Long non-coding RNAs: new players in cell differentiation and development. *Nat. Rev. Genet.* 15, 7–21. doi: 10.1038/nrg3606
- Ghanbari, M., Ikram, M. A., de Looper, H. W. J., Hofman, A., Erkeland, S. J., Franco, O. H., et al. (2016). Genome-wide identification of microRNA-related variants associated with risk of Alzheimer's disease. *Sci. Rep.* 6:28387. doi: 10.1038/srep28387
- Grupe, A., Li, Y., Rowland, C., Nowotny, P., Hinrichs, A. L., Smemo, S., et al. (2006). A scan of chromosome 10 identifies a novel locus showing strong association with late-onset Alzheimer disease. *Am. J. Hum. Genet.* 78, 78–88. doi: 10.1086/498851
- GTEX (2013). The genotype-tissue expression (GTEx) project. *Nat. Genet.* 45, 580–585. doi: 10.1038/ng.2653
- Han, S., Zhuang, H., Lee, P. Y., Li, M., Yang, L., Nigrovic, P. A., et al. (2019). Differential responsiveness of monocyte and macrophage subsets to interferon. *Arthritis Rheumatol.* 72, 100–113. doi: 10.1002/art.41072
- Harper, S. (2014). Economic and social implications of aging societies. *Science* 346, 587–591. doi: 10.1126/science.1254405
- Hensman Moss, D. J., Flower, M. D., Lo, K. K., Miller, J. R., van Ommen, G. B., t Hoen, P. A., et al. (2017). Huntington's disease blood and brain show a common gene expression pattern and share an immune signature with Alzheimer's disease. *Sci. Rep.* 7:44849. doi: 10.1038/srep44849
- Huang, Z., Shi, J., Gao, Y., Cui, C., Zhang, S., Li, J., et al. (2019). HMDD v3.0: a database for experimentally supported human microRNA-disease associations. *Nucleic Acids Res.* 47, D1013–D1017. doi: 10.1093/nar/gky1010
- Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S. L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. *Genome Biol.* 14:R36. doi: 10.1186/gb-2013-14-4-r36
- Kinnersley, B., Labussiere, M., Holroyd, A., Di Stefano, A. L., Broderick, P., Vijaykrishnan, J., et al. (2015). Genome-wide association study identifies multiple susceptibility loci for glioma. *Nat. Commun.* 6:8559. doi: 10.1038/ncomms9559
- Kold-Christensen, R., and Johannsen, F. (2019). Methylglyoxal metabolism and aging-related disease: moving from correlation toward causation. *Trends Endocrinol. Metab.* 31, 81–92. doi: 10.1016/j.tem.2019.10.003
- Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559
- Li, J. H., Liu, S., Zhou, H., Qu, L. H., and Yang, J. H. (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–D97. doi: 10.1093/nar/gkt1248
- Liang, W. S., Dunckley, T., Beach, T. G., Grover, A., Mastroeni, D., Walker, D. G., et al. (2007). Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. *Physiol. Genomics* 28, 311–322. doi: 10.1152/physiolgenomics.00208.2006
- Mercer, T. R., and Mattick, J. S. (2013). Structure and function of long noncoding RNAs in epigenetic regulation. *Nat. Struct. Mol. Biol.* 20, 300–307. doi: 10.1038/nsmb.2480
- Nt, P., Ds, M., Mm, N., In, M., and Ti, V. (2018). Investigation of circulating serum MicroRNA-328-3p and MicroRNA-3135a expression as promising novel biomarkers for autism spectrum disorder. *Balkan J. Med. Genet.* 21, 5–12. doi: 10.2478/bjmg-2018-0026
- Papaio, R., Kunderfranco, P., Stirparo, G. G., Latronico, M. V., and Condorelli, G. (2013). Long noncoding RNA: a new player of heart failure? *J. Cardiovasc. Transl. Res.* 6, 876–883. doi: 10.1007/s12265-013-9488-6
- Peters, M. J., Joehanes, R., Pilling, L. C., Schurmann, C., Conneely, K. N., Powell, J., et al. (2015). The transcriptional landscape of age in human peripheral blood. *Nat. Commun.* 6:8570. doi: 10.1038/ncomms9570
- Pinero, J., Bravo, A., Queralt-Rosinach, N., Gutierrez-Sacristan, A., Deu-Pons, J., Centeno, E., et al. (2017). DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic Acids Res.* 45, D833–D839. doi: 10.1093/nar/gkw943
- Prasad, K., Abdel-Hameed, A. A. E., Xing, D., and Reddy, A. S. N. (2016). Global gene expression analysis using RNA-seq uncovered a new role for SRI/CAMTA3 transcription factor in salt stress. *Sci. Rep.* 6:27021. doi: 10.1038/srep27021
- Provost, P. (2010). Interpretation and applicability of microRNA data to the context of Alzheimer's and age-related diseases. *Aging* 2, 166–169. doi: 10.18632/aging.100131
- Puthiyedth, N., Riveros, C., Berretta, R., and Moscato, P. (2016). Identification of differentially expressed genes through integrated study of Alzheimer's Disease affected brain regions. *PLoS One* 11:e0152342. doi: 10.1371/journal.pone.0152342
- Rinn, J. L., and Chang, H. Y. (2012). Genome regulation by long noncoding RNAs. *Annu. Rev. Biochem.* 81, 145–166. doi: 10.1146/annurev-biochem-051410-092902
- Sadlon, A., Takousis, P., Alexopoulos, P., Evangelou, E., Prokopenko, I., and Pernecky, R. (2019). miRNAs identify shared pathways in Alzheimer's and Parkinson's Diseases. *Trends Mol. Med.* 25, 662–672. doi: 10.1016/j.molmed.2019.05.006
- Scheckel, C., Drapeau, E., Frias, M. A., Park, C. Y., Fak, J., Zucker-Scharff, I., et al. (2016). Regulatory consequences of neuronal ELAV-like protein binding to coding and non-coding RNAs in human brain. *Elife* 5:e10421. doi: 10.7554/eLife.10421
- Shi, X., Sun, M., Liu, H., Yao, Y., and Song, Y. (2013). Long non-coding RNAs: a new frontier in the study of human diseases. *Cancer Lett.* 339, 159–166. doi: 10.1016/j.canlet.2013.06.013
- Sood, S., Gallagher, I. J., Lunnon, K., Rullman, E., Keohane, A., Crossland, H., et al. (2015). A novel multi-tissue RNA diagnostic of healthy ageing relates to cognitive health status. *Genome Biol.* 16:185. doi: 10.1186/s13059-015-0750-x
- Sultan, M., Schulz, M. H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., et al. (2008). A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. *Science* 321, 956–960. doi: 10.1126/science.1160342
- Sultmann, H., and Diederichs, S. (2014). Long noncoding RNA: “LNCs” to cancer. *Eur. Urol.* 65, 1152–1153. doi: 10.1016/j.eururo.2014.01.032
- Sun, J., Shi, H., Wang, Z., Zhang, C., Liu, L., Wang, L., et al. (2014). Inferring novel lncRNA-disease associations based on a random walk model of a lncRNA functional similarity network. *Mol. Biosyst.* 10, 2074–2081. doi: 10.1039/c3mb70608g
- Tacutu, R., Thornton, D., Johnson, E., Budovsky, A., Barardo, D., Craig, T., et al. (2018). Human ageing genomic resources: new and updated databases. *Nucleic Acids Res.* 46, D1083–D1090. doi: 10.1093/nar/gkx1042
- Toledo, J. B., Zetterberg, H., van Harten, A. C., Glodzik, L., Martinez-Lage, P., Bocchio-Chiavetto, L., et al. (2015). Alzheimer's disease cerebrospinal fluid biomarker in cognitively normal subjects. *Brain* 138(Pt 9), 2701–2715. doi: 10.1093/brain/awv199

- Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016
- Twine, N. A., Janitz, K., Wilkins, M. R., and Janitz, M. (2011). Whole transcriptome sequencing reveals gene expression and splicing differences in brain regions affected by Alzheimer's disease. *PLoS One* 6:e16266. doi: 10.1371/journal.pone.0016266
- Uberti, D., Cenini, G., Bonini, S. A., Barcikowska, M., Styczynska, M., Szybinska, A., et al. (2010). Increased CD44 gene expression in lymphocytes derived from Alzheimer disease patients. *Neurodegener. Dis.* 7, 143–147. doi: 10.1159/000289225
- Wan, G., Zhou, W., Hu, Y., Ma, R., Jin, S., Liu, G., et al. (2016). Transcriptional regulation of lncRNA genes by histone modification in Alzheimer's Disease. *Biomed. Res. Int.* 2016:3164238. doi: 10.1155/2016/3164238
- Xiong, J., Lu, X., Zhou, Z., Chang, Y., Yuan, D., Tian, M., et al. (2012). Transcriptome analysis of the model protozoan, *Tetrahymena thermophila*, using Deep RNA sequencing. *PLoS One* 7:e30630. doi: 10.1371/journal.pone.0030630
- Xu, M., Zhang, D. F., Luo, R., Wu, Y., Zhou, H., Kong, L. L., et al. (2018). A systematic integrated analysis of brain expression profiles reveals YAP1 and other prioritized hub genes as important upstream regulators in Alzheimer's disease. *Alzheimers Dement* 14, 215–229. doi: 10.1016/j.jalz.2017.08.012
- Zhang, D. F., Fan, Y., Xu, M., Wang, G., Wang, D., Li, J., et al. (2019). Complement C7 is a novel risk gene for Alzheimer's disease in Han Chinese. *Natl. Sci. Rev.* 6, 257–274. doi: 10.1093/nsr/nwy127
- Zhou, M., Zhao, H., Wang, X., Sun, J., and Su, J. (2019). Analysis of long noncoding RNAs highlights region-specific altered expression patterns and diagnostic roles in Alzheimer's disease. *Brief Bioinform* 20, 598–608. doi: 10.1093/bib/bby021
- Zhou, M., Zhao, H., Xu, W., Bao, S., Cheng, L., and Sun, J. (2017). Discovery and validation of immune-associated long non-coding RNA biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell lymphoma. *Mol. Cancer* 16:16. doi: 10.1186/s12943-017-0580-4

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Shi, Liu, Yang, Xu, Cai, Wang, Zhao, Shao and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.