BIOINFORMATICS ORIGINAL PAPER

Vol. 29 no. 20 2013, pages 2596–2602
doi:10.1093/bioinformatics/btt423

Gene expression

Advance Access publication August 29, 2013

Identification of active transcription factor and miRNA regulatory pathways in Alzheimer's disease

Wei Jiang[1,†], Yan Zhang[1,†], Fanlin Meng[1], Baofeng Lian[1], Xiaowen Chen[1], Xuexin Yu[1], Enyu Dai[1], Shuyuan Wang[1], Xinyi Liu[1], Xiang Li[2], Lihong Wang[3,*] and Xia Li[1,*]

[1]College of Bioinformatics Science and Technology, [2]Department of Internal Neurology, Fourth Affiliated Hospital and [3]Institute of Cancer Prevention and Treatment, Harbin Medical University, Harbin 150081, China

Associate Editor: Gunnar Ratsch

ABSTRACT

Motivation: Alzheimer's disease (AD) is a severe neurodegenerative disease of the central nervous system that may be caused by perturbation of regulatory pathways rather than the dysfunction of a single gene. However, the pathology of AD has yet to be fully elucidated.

Results: In this study, we systematically analyzed AD-related mRNA and miRNA expression profiles as well as curated transcription factor (TF) and miRNA regulation to identify active TF and miRNA regulatory pathways in AD. By mapping differentially expressed genes and miRNAs to the curated TF and miRNA regulatory network as active seed nodes, we obtained a potential active subnetwork in AD. Next, by using the breadth-first-search technique, potential active regulatory pathways, which are the regulatory cascade of TFs, miRNAs and their target genes, were identified. Finally, based on the known AD-related genes and miRNAs, the hypergeometric test was used to identify active pathways in AD. As a result, nine pathways were found to be significantly activated in AD. A comprehensive literature review revealed that eight out of nine genes and miRNAs in these active pathways were associated with AD. In addition, we inferred that the pathway hsa-miR-146a→STAT1→MYC, which is the source of all nine significantly active pathways, may play an important role in AD progression, which should be further validated by biological experiments. Thus, this study provides an effective approach to finding active TF and miRNA regulatory pathways in AD and can be easily applied to other complex diseases.

Contact: lixia@hrbmu.edu.cn or lw2247@gmail.com.

Supplementary information: Supplementary data are available at *Bioinformatics* online.

Received on November 22, 2012; revised on July 17, 2013; accepted on July 18, 2013

1 INTRODUCTION

In the developed countries, Alzheimer's disease (AD) is the sixth leading cause of all deaths. Deaths attributable to AD have been increasing dramatically, whereas other major causes of death have been decreasing (Thies and Bleiler, 2011). In addition, AD is one of the most costly diseases for society (Holt *et al.*, 2009; Thies and Bleiler, 2011). However, the molecular mechanism of AD is not fully clear. It is thought that many factors and their interactions contribute to the pathogenesis of AD (Ikonen *et al.*, 2003; Lahiri *et al.*, 2004; Xia *et al.*, 1997). Thus, advanced research regarding the mechanism of AD is of great importance.

With the emergence of molecular networks, such as protein–protein interaction networks (PPINs) and transcription regulatory networks, many studies initially focus on topological properties when analyzing network organization, architecture or evolution. However, most existing studies are based on PPINs and transcription regulatory networks and do not integrate important post-transcriptional regulation.

miRNA is a type of short non-coding RNA that participates in post-transcriptional gene regulation. By binding to target mRNAs with partially complementary sequences, it causes translational repression or target degradation (Bartel, 2004, 2009). Studies have increasingly shown that miRNA is closely associated with the onset and development of complex human diseases (Li *et al.*, 2012), such as cancer (Meltzer, 2005), diabetes (Kantharidis *et al.*, 2011) and neurodegenerative disease (Junn and Mouradian, 2012). With increasing research on AD, many studies have shown that miRNAs are important players in the development of AD. For example, Wang *et al.* found that miR-107 expression levels decreased significantly, even in patients who were in the earliest stages of pathology. Wang *et al.* (2008) further indicated that miR-107 may cause AD through the regulation of BACE1. A recent study by Shioya *et al.* (2010) found that the underexpression of miR-29a affected neurodegenerative processes by enhancing neuronal NAV3 expression in AD brains. In addition, Hebert *et al.* (2008) characterized a miRNA cluster miR-29a/b-1 that was significantly and specifically downregulated in AD patients, and found that the loss of miRNA increases BACE1 and Aβ expression levels in AD. These findings suggest that aberrant miRNA expression is strongly related to AD. In addition, miRNA and transcription factor (TF) do not function in isolation. Chen *et al.* (2011) found that the coordinated regulation of TF and miRNA may be involved in various biological processes, and the disruption of this coordination may lead to cancer. Thus, integrated analysis of transcriptional and post-transcriptional regulation could provide a comprehensive regulatory map for the etiological study of complex diseases.

In recent years, gene expression profiles have been taken into account during network analysis to identify the underlying mechanisms of complex diseases. Ruan *et al.* developed a general

*To whom correspondence should be addressed.

†The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

coexpression network-based approach to discovering network modular structures that was effectively applied to the study of the pathological mechanisms underlying lymphoma (Cho *et al.*, 2011; Ruan *et al.*, 2010). Vaske *et al.* (2010) presented a method for predicting the activities of curated pathways from the Pathway Interaction Database in glioblastoma multiform and breast cancer by considering gene coordinate expression and the correlation between copy number and gene expression in the same pathway. In addition, much effort has been devoted to identifying the active subnetworks or modules in diseases from large molecular networks. For example, Ideker *et al.* (2002) proposed a method for screening a molecular interaction network that contained protein–protein and protein–DNA interactions to find active subnetworks. A scoring system was implemented to capture the degree of change in gene expression. A search algorithm based on simulated annealing was used to identify the subnetwork with the highest score. Gene expression differences between normal and disease samples were interpreted as the degree of gene activation. Backes *et al.* (2012) exploited the same 'active' concept to find the deregulated subnetwork in the regulatory network of KEGG using an integer linear programming approach. Gaire *et al.* (2013) extracted the active subnetwork in the network that consists of protein interaction network and gene regulatory network by applying a mixed integer programming model. However, finding the subnetwork or module that is significantly related to a disease state is still challenging. Compared with coexpression networks, curated pathways and complex subnetworks, the regulatory pathway, which presents a cascade of regulators and target genes, is easily elucidated and validated by biologists. From disease-related pathways, we may find key factors that are located upstream of the pathway and that participate in multiple pathways. The activation of key factors may be the cause of disease onset. Thus, identifying active regulatory pathways is crucial for dissecting the pathology of complex diseases. Keller *et al.* (2009) presented an algorithm for detecting differentially regulated paths (pathways) from a given biological network based on gene set enrichment analysis. However, they did not consider genes whose expression level were not detected by microarray technology and restricted the length of paths (pathways).

In this study, we identified active transcriptional and post-transcriptional regulatory pathways in AD based on AD-related mRNA and miRNA expression profiles as well as miRNA and TF regulation. To obtain stable AD signatures [differentially expressed (DE) genes], we used a meta-analysis to analyze multiple AD-related gene expression profiles. In addition, a breadth-first-search (BFS) approach was applied to find potential active TF-miRNA regulatory pathways that may contain non-DE genes. Finally, the significance of all of the potential pathways was evaluated by a hypergeometric test based on known AD-related genes and miRNAs. The identified pathway had a simple network structure and an easy interpretation that could feasibly be validated by biological experiments. Our study thus provides a novel insight into the causes and mechanisms of AD.

2 MATERIALS AND METHODS

We proposed a novel approach to identify active TF-miRNA regulatory pathways in AD. Firstly, we detected consistent differentially expressed genes (CDEGs) by applying a meta-analysis to multiple gene microarray datasets. The differentially expressed miRNAs (DEmiRs) were derived from one miRNA expression profile. Secondly, by mapping the CDEGs and DEmiRs to the curated TF-miRNA regulatory network as active seed nodes and connecting the active seed nodes with their immediate neighbors, we obtained the potential active TF-miRNA regulatory subnetwork in AD. Thirdly, by using a BFS algorithm, we identified all of the directed acyclic paths between 0-indegree nodes, where the indegree is 0, and 0-outdegree nodes, where the outdegree is 0, which were defined as potential active TF-miRNA regulatory pathways in AD. Finally, known AD-associated genes and miRNAs were mapped to potential active pathways. Through the hypergeometric test, we identified the active TF-miRNA regulatory pathways that were significantly related to AD. An overview of the approach is shown in Figure 1.

2.1 AD-related mRNA expression profiles

AD is a complex disease, and different brain regions exhibit diverse gene expression patterns (Liang *et al.*, 2007). However, several studies have suggested that some important common features may be shared among different AD brain regions. Liu *et al.* (2010) detected crosstalk and dysfunction in AD-related pathways in multiple brain regions of AD patients and identified common dysfunctions. Using a microarray dataset on six AD brain regions (GSE5281), Liang *et al.* (2012) identified a significantly perturbed subnetwork in each brain region and found that these perturbed subnetworks significantly overlapped with each other. In this study, to obtain stable and consistent AD gene signatures, we downloaded 4 AD-related mRNA expression profiles (GSE16759, GSE12685, GSE1297 and GSE5281) from the Gene Expression Omnibus (GEO) database (Edgar *et al.*, 2002). All of the data were normalized and log 2 transformed (details in Supplementary file S1). For each expression profile, probe sets were mapped to Entrez Gene IDs. If multiple probe sets corresponded to the same gene, then the expression values of these probe sets were averaged. Genes that appeared in all of the expression profiles were considered in the analysis. The microarray datasets comprise nine case-control studies (Supplementary File S1 and S2) according to different experiments and brain regions, which were further analyzed by the meta-analysis.

2.2 DEmiRs in AD

Because high-throughput techniques have thus far rarely been used to investigate miRNA expression in AD, we only obtained one miRNA

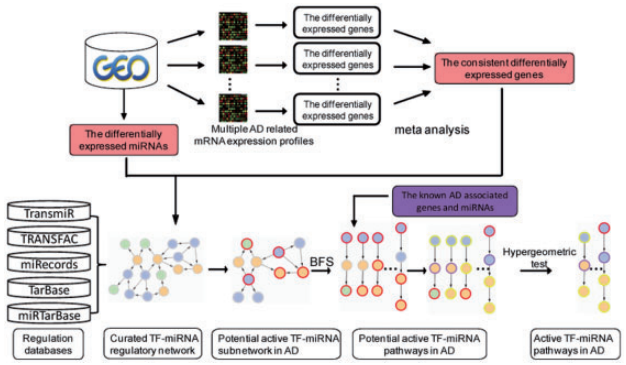

Fig. 1. Flow chart of the proposed approach. The orange nodes represent miRNAs, the blue nodes represent TFs and the green nodes represent target genes. The nodes with red border indicate the DE genes/miRNAs, the nodes with purple border indicate the known AD-associated genes/miRNAs and the nodes with yellow border indicate the known and DE genes/miRNAs

microarray dataset from GEO (GSE16759), in which Nunez-Iglesias *et al.* (2010) detected both miRNA and mRNA expression in AD patients and controls. They identified 48 DEmiRs at the false discovery rate (FDR) level of 0.05 by the empirical Bayes procedure.

2.3 Curated TF and miRNA regulatory network

The TF and miRNA regulatory network was constructed by integrating five curated data resources: TRANSFAC (Wingender *et al.*, 2000), TransmiR (Wang *et al.*, 2010), miRTarBase (Hsu *et al.*, 2011), miRecords (Xiao *et al.*, 2009) and TarBase (Sethupathy *et al.*, 2006). The curated human TF-gene regulations were derived from the TRANSFAC database (version 11.4). The curated human TF-miRNA regulations were obtained from the TransmiR database (version 1.2). The curated human miRNA-gene regulations were obtained from the union of miRecords (version 3), TarBase (version 5.0) (excluding the FALSE support type of the miRNA-gene interactions) and miRTarBase databases (release 2.5). All curated regulations were supplied in the Supplementary File S3. Within the curated regulatory network, all of the redundant edges were collapsed into a single edge, and all of the self-directed edges were pruned from the network.

2.4 Known AD-associated genes and miRNAs

The known AD-associated genes were derived from the GeneCards database (Safran *et al.*, 2002). The disease genes presented in GeneCards were extracted from multiple databases that included known disease-associated genes. The known AD-associated miRNAs were derived from HMDD (Lu *et al.*, 2008) and the miR2Disease database (Jiang *et al.*, 2009). Both of these databases manually collected the associations between miRNAs and diseases from published studies.

2.5 Meta-analysis

Different gene expression profiles may produce different DE genes. To obtain stable gene signatures, we used a meta-analysis (Ramasamy *et al.*, 2008) to identify CDEGs through a combination of nine case-control studies. The detail description of meta-analysis was provided in Supplementary File S1. In this study, we used the SMVar (Jaffrezic *et al.*, 2007) method to detect DE genes in each case-control study. The meta-analysis was implemented by using the R package metaMA (Marot *et al.*, 2009).

2.6 Identification of a potential active TF-miRNA regulatory subnetwork in AD

Previous studies have shown that disease genes do not always show differential expression in microarray experiments, which indicates that some disease-related key genes could lurk among non-DE genes (Nitsch *et al.*, 2009; Zhao *et al.*, 2011). As a result, we hypothesized that the CDEGs, DEmiRs and their immediate neighbors in the curated TF-miRNA regulatory network potentially contributed to the pathology of AD. We mapped the CDEGs and DEmiRs into the regulatory network as active seed nodes and connected them with their neighbors to produce the potential active TF-miRNA regulatory subnetwork.

2.7 Identification of potential active TF-miRNA regulatory pathways in AD

A subnetwork often has a complex structure, even if it is distilled from the original background network. These complex connections impede the interpretation and validation of the subnetwork of interest. In this study, we focused on regulatory pathways, which were the paths connected to multiple TFs, miRNAs and target genes in the curated TF-miRNA regulatory network. Identifying the active regulatory pathways in AD not only uncovered transcriptional and post-transcriptional regulatory cascades but also shed light on the molecular mechanisms of AD.

From the potential active TF-miRNA regulatory subnetwork, we identified all directed acyclic paths from 0-indegree nodes to 0-outdegree nodes. The gene/miRNA with a 0-indegree cannot be regulated by other regulators, which indicates that it is located upstream of the regulatory pathway. Similarly, the gene/miRNA with a 0-outdegree does not regulate other genes/miRNAs, which means that it is located downstream of the regulatory pathway. The upstream genes/miRNAs are important because their activation could cause a cascade effect that results in the alteration of downstream gene/miRNA expression and leads to AD. Thus, by searching all of the pathways/paths between 0-indegree genes/miRNAs and 0-outdegree genes/miRNAs, we could find key upstream genes/miRNAs in the regulatory pathways. To accomplish this task, the potential active subnetwork was treated as a directed graph. Firstly, we used BFS algorithm to traverse all vertexes in the graph. Secondly, based on the results of graph traversal, backtracking method was used to extract all paths from 0-indegree nodes to 0-outdegree nodes. The pseudo-code for finding the pathways is described in Supplementary File S1.

For the example graph in Figure 2, we can identify all of the directed acyclic paths between the 0-indegree node and the 0-outdegree node. They are $1\rightarrow2\rightarrow4\rightarrow7$, $1\rightarrow2\rightarrow4\rightarrow8$, $1\rightarrow2\rightarrow5\rightarrow8$ and $1\rightarrow3\rightarrow6\rightarrow9$.

In this study, the directed acyclic paths with more than 2 nodes were considered to be potential active regulatory pathways, in which there was at least one DE node and no more than one non-DE node or node without expression values between the two DE nodes.

2.8 Evaluation of potential active TF-miRNA regulatory pathways in AD

Here, we defined a coverage rate (*CR*) of known AD-associated genes and miRNAs in the potential active pathway to measure the strength of the relationships between the potential active pathway and AD.

CR was calculated as

$$CR = \frac{N_D}{N_T}$$

where N_D represents the number of known AD-associated genes and miRNAs in the pathway, and N_T represents the total number of genes and miRNAs in the pathway.

Next, we used the hypergeometric test to evaluate the statistical significance of the *CR* value. A low *P*-value indicates that the observed *CR* value is unlikely to occur by chance and the pathway exhibits a greater than expected trend toward participating in AD.

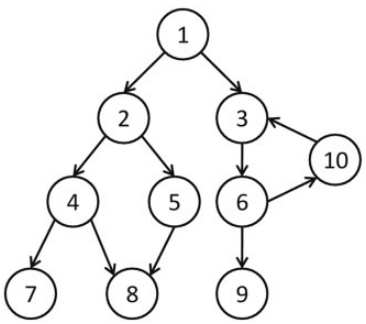

Fig. 2. An example of how to find the pathways

3 RESULTS

3.1 DE genes and miRNAs

By using a meta-analysis to combine nine case-control microarray studies (Supplementary Table S1), we identified 1872 CDEGs at the FDR level of 0.01.

Nunez-Iglesias *et al.* obtained 48 DEmiRs by applying the empirical Bayes method to miRNA expression profiles in AD patients and normal controls (Nunez-Iglesias *et al.*, 2010). In the present study, 29 DEmiRs that were recorded in the miRBase database were used.

3.2 Potential active TF-miRNA regulatory subnetwork in AD

By integrating five databases (TRANSFAC, TransmiR, miRTarBase, miRecords and TarBase) that address transcriptional and post-transcriptional regulations, we constructed the curated TF-miRNA regulatory network, which included 411 TFs, 387 miRNAs, 2300 target genes and 6036 regulations (Fig. 3A). We mapped the 1872 CDEGs and 29 DEmiRs to the curated TF-miRNA regulatory network and set them to be active seeds. Next, we constructed the potential active TF-miRNA regulatory subnetwork by connecting all of the active seeds with their immediate neighbors (Fig. 3B). The subnetwork comprised 127 TFs, 195 miRNAs, 649 target genes, and 1206 edges, in which 343 genes and 16 miRNAs were DE.

3.3 The active TF-miRNA regulatory pathways in AD

We used the BFS approach (see 'Materials and Methods' section) to find all of the directed acyclic paths from 0-indegree nodes to 0-outdegree nodes in the potential active TF-miRNA regulatory subnetwork. As a result, 14 644 paths with more than two nodes were obtained, and these paths were regarded as the potential active TF-miRNA regulatory pathways in AD, which comprised 641 genes and miRNAs. The length of all of the potential active pathways ranged from 3 to 16, and the average was 6.99.

In addition, we derived 27 known AD-associated genes from the GeneCards database (Safran *et al.*, 2002) and 45 known AD-associated miRNAs from the HMDD database (Lu *et al.*, 2008) and miR2Disease database (Jiang *et al.*, 2009) to evaluate the significance of the identified potential active pathways. There were 29 AD-associated genes and miRNAs mapped in the potential active pathways. The *CR* of the known AD-associated genes and miRNAs of the potential active pathway was used to measure the strength of the association between the potential active pathway and AD. Next, we identified the significantly active pathways using a hypergeometric test (see 'Materials and Methods' section). The potential active pathways with $P<0.0005$ were considered to be the active TF-miRNA regulatory pathways in AD. As a result, we identified nine active pathways. Moreover, we further adjusted *P*-values for multiple testing using FDR and found that the FDR of the nine pathways are less than 0.2 (Table 1). The union of the nine active pathways is visualized in Figure 4, and identical nodes of the active pathways were merged.

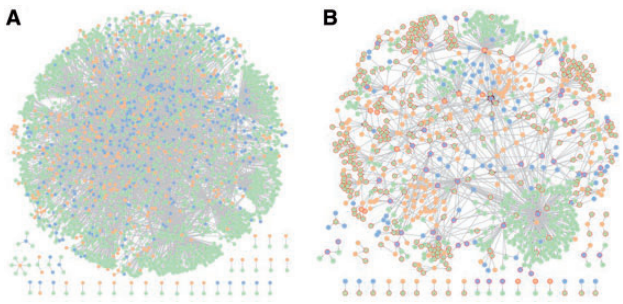

Fig. 3. The curated TF-miRNA regulatory network and potential active subnetwork in AD. The orange nodes represent miRNAs, the blue nodes represent TFs and the green nodes represent target genes. (**A**) The curated TF-miRNA regulatory network; (**B**) The potential active TF-miRNA regulatory subnetwork in AD. The red border indicates the DE genes and miRNAs

We validated the active pathways from three aspects. Firstly, we tested the extent to which the elements in all of the active pathways appeared in the set of known AD-associated genes and miRNAs using the hypergeometric test. There were 29 known AD-associated genes and miRNAs in the potential active regulatory pathways, which comprised 641 genes and miRNAs. And there were 18 genes and miRNAs in all of the active pathways, which included nine known AD-associated genes and miRNAs. As a result, the elements in all of the active pathways significantly enriched the set of known AD-associated genes and miRNAs ($P=7.88\times10^{-9}$). However, this result may not be surprising because the known AD-associated genes and miRNAs have been used to investigate the significance of potential active pathways. Thus, we further evaluated the active pathways through the curated AD pathway (hsa05010) from KEGG (release 65.0). We applied miRPath v2.0 (Vlachos *et al.*, 2012) to find miRNAs that significantly regulate the hsa05010 pathway ($P<0.01$). Altogether, all potential active regulatory pathways and all active regulatory pathways include 17 and 4 genes and miRNAs that are associated with hsa05010 pathway, respectively. As a result, the hypergeometric test *P*-value is 8.29×10^{-4}, which indicates that the genes and miRNAs in our identified active pathways significantly associated with the known AD pathway. Finally, through a literature review, we found that eight out of the nine remaining genes and miRNAs are closely related to AD (Table 2). Here, we could not find any direct evidence to support an important role of protein kinase C alpha (PRKCA) in AD. However, Barton *et al.* (2004) showed that the PRKCA gene was associated with multiple sclerosis, which is a common inflammatory disease of the central nervous system and presents some similarities to AD. Our results indicate that PRKCA might exert an effect through hsa-miR-146a regulation in AD. Additionally, we also found some regulatory interactions in active pathways were relevant to AD. For example, two regulations of SP1→APOE and SP1→MPO appear in the active pathways. The polymorphisms in the promoter regions of the two target genes influence the SP1 regulations in AD. Maloney *et al.* (2010) demonstrated that two SNPs (A-491T and G-219T) in APOE promoter region have significant association with instance of AD. They also proposed that SP1 is the candidates for regulatory control of the two polymorphic sites. Leininger-Muller *et al.* introduced that the G allele of G-463A MPO polymorphism is associated with a higher level of MPO

Table 1. Active TF-miRNA regulatory pathways in AD

Active TF-miRNA regulatory pathway	Number of known AD genes and miRNAs	Pathway length	*CR* value	*P*-value	FDR
hsa-miR-146a → STAT1 → MYC → hsa-miR-15a → NFKB1 → hsa-miR-29b → SP1 → MPO	4	8	0.50	0.00021	0.11021
hsa-miR-146a → STAT1 → MYC → hsa-miR-15a → NFKB1 → hsa-miR-29b → SP1 → APOE	4	8	0.50	0.00021	0.11021
hsa-miR-146a → STAT1 → MYC → hsa-miR-15a → NFKB1 → hsa-miR-29b → SP1 → NOS3	4	8	0.50	0.00021	0.11021
hsa-miR-146a → STAT1 → MYC → hsa-miR-29b → SP1 → PRKCA → hsa-miR-15a → APP	4	8	0.50	0.00021	0.11021
hsa-miR-146a → STAT1 → MYC → hsa-miR-29b → SP1 → E2F1 → hsa-miR-15a → APP	4	8	0.50	0.00021	0.11021
hsa-miR-146a → STAT1 → MYC → hsa-miR-15a → NFKB1 → hsa-miR-29b → SP1 → PSEN1	4	8	0.5	0.00021	0.11021
hsa-miR-146a → STAT1 → MYC → hsa-miR-15a → NFKB1 → hsa-miR-29b → SP1 → hsa-miR-34c → NOTCH3	4	9	0.44	0.000366	0.11021
hsa-miR-146a → STAT1 → MYC → hsa-miR-15a → NFKB1 → hsa-miR-29b → SP1 → hsa-miR-34c → NOTCH2	4	9	0.44	0.000366	0.11021
hsa-miR-146a → STAT1 → MYC → hsa-miR-15a → NFKB1 → hsa-miR-29b → SP1 → hsa-miR-34c → NOTCH1	4	9	0.44	0.000366	0.11021

and could be overrepresented in AD subjects (Leininger-Muller *et al.*, 2003). Previous study also demonstrated that the G-463A is in the promoter region of MPO, and the A allele rather than G leads to a decreased MPO expression by destroying a SP1 binding site (Piedrafita *et al.*, 1996). Aβ is a peptide of 36–43 amino acids that is processed from the amyloid precursor protein (APP) and is expressed in many tissues and concentrated in neuronal synapses (Krishnappa, 2011). It is commonly known that the accumulation of Aβ peptides could lead to the formation of

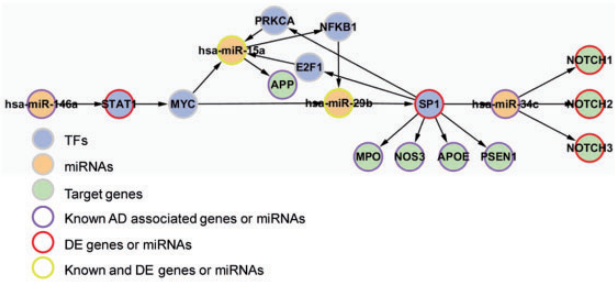

Fig. 4. Union of nine active TF-miRNA regulatory pathways in AD

amyloid plaques in the brain parenchyma, which is considered to be a key step in the pathogenesis of AD (Ramberg *et al.*, 2011). Therefore, understanding the processing of APP is crucial for development of AD therapeutics. However, little is known about the mechanisms that contribute to Aβ accumulation in sporadic AD. Here, we identified two active TF-miRNA regulatory pathways in AD that included APP (Table 1 and Fig. 4). Proteolytic processing of APP at the β site is essential for generating Aβ. BACE1, which is the major β-secretase involved in cleaving APP, has been identified as a type 1 membrane-associated aspartyl protease. Christensen *et al.* (2004) found that the TF SP1 plays an important role in processing APP to generate Aβ in Alzheimer's disease through regulation of BACE1. miRNAs can regulate gene expression at the post-transcriptional level. Hebert *et al.* (2008) investigated miRNA expression changes in AD patients and found that a loss of hsa-miR-29b could increase the BACE1 expression level, which indicated a potential causal relationship between the expression of hsa-miR-29b and the accumulation of Aβ. As mentioned above, all of these factors affect APP processing by targeting BACE1. Hebert *et al.* also found that hsa-miR-15a were significantly altered in AD brain and predicted that has-miR-15a regulates APP (Hebert *et al.*, 2008). This regulation of hsa-miR-15a→APP appeared in the active pathways. In this study, our results suggest an alternative possibility for the regulation of Aβ levels through cascade regulatory pathways.

Our results confirmed that the disease-related genes and miRNAs did not always present aberrant expression and that non-DE genes and miRNAs may be key players in the disease. In addition, our method can also be used to find key regulators for which there is no expression information.

Considerable evidence supports the conclusion that neuroinflammation is associated with AD pathology (Tuppo and Arias, 2005). Lukiw *et al.* (2008) suggested that NFKB-sensitive miR-146a-mediated modulation of the expression of the CFH gene, which is an important repressor of the inflammatory response of the brain, may have an impact on AD. Kitamura *et al.* (1997) indicated that increased expression of NFKB and STAT1 in cell nuclei may be involved in inflammatory activation in AD brains.

Ferrer and Blanco (2000) detected the expression of MYC protein by using western blot and single and double-labeling immunohistochemistry in AD, and found that increased expression of MYC in reactive astrocytes most likely plays a role in reactive astrocytosis in human neurodegenerative disorders.

Table 1 and Figure 4 show that hsa-miR-146a→ STAT1 →MYC is located upstream of the regulatory cascade and is

Table 2. Genes and miRNAs excluding known AD elements in the nine active pathways

Genes	Description	Literature support
E2F1	E2F transcription factor 1	Chen *et al.* (2003)
MYC	v-myc myelocytomatosis viral oncogene homolog (avian)	Ferrer and Blanco (2000)
NFKB1	Nuclear factor of kappa light polypeptide gene enhancer in B-cells 1	Lukiw and Alexandrov (2012); Lukiw *et al.* (2008)
NOTCH1	Notch 1	Berezovska *et al.* (1999)
NOTCH2	Notch 2	Chen *et al.* (2006)
NOTCH3	Notch 3	Sisodia and St George-Hyslop (2002)
PRKCA	Protein kinase C, alpha	
SP1	Sp1 transcription factor	Citron *et al.* (2008) Santpere *et al.* (2006)
STAT1	Signal transducer and activator of transcription 1, 91 kDa	Kitamura *et al.* (1997)

the source of nine active pathways. The activation of this source could cause the emergence of a cascade effect and result in disordered expression of downstream genes and miRNA. Thus, hsa-miR-146a may be a potential drug target for disease therapy, and the pathway hsa-miR-146a→STAT1→MYC may have a crucial role in AD that should be further validated by biological experiments.

4 DISCUSSION

The dysregulation of miRNAs, a type of key post-transcriptional regulators, has been found in many diseases, such as cancers and AD. In this study, we proposed a novel approach to identify active TF-miRNA regulatory pathways by integrating AD-related mRNA and miRNA expression profiles and transcriptional and post-transcriptional regulation. The pathways we identified could help to provide biological insights in AD and could feasibly be validated by biological experiments. We defined the DE genes and miRNAs as active seed nodes in the curated TF-miRNA regulatory network. To incorporate regulators without aberrant expression or expression values, we constructed a potential active subnetwork by connecting the active seed nodes with their immediate neighbors. Next, the BFS method was used to find the potential active regulatory pathways. Finally, based on the known AD-related genes and miRNAs, we used the hypergeometric test to evaluate the significance of the associations between potential active pathways and AD. As a result, we identified nine active TF-miRNA regulatory pathways that were significantly related to AD. All of the regulators and target genes in these active pathways had direct literature support, except for PRKCA. We also found two active pathways that may contribute to the accumulation of Aβ peptides and the formation of amyloid plaques in the brain parenchyma, which are key steps in the pathogenesis of AD. In addition, the pathway hsa-miR-146a→STAT1→MYC appeared in all nine active pathways and was the source of the cascade regulation. Thus, we inferred that the pathway hsa-miR-146a→STAT1→MYC may have a crucial role in AD progression, which should be further validated by biological experiments.

Our study provides an effective approach to identifying active TF-miRNA regulatory pathways and a novel insight into the pathogenesis and development of AD. However, the regulation between miRNA and mRNA is complex, as suggested by the miRNA sponge model, which is a new type of miRNA-mediated post-transcriptional regulation (Sumazin *et al.*, 2011). Many studies have suggested that miRNA sponges play important roles in the occurrence and development of diseases, such as autoimmune inflammation (Zhu *et al.*, 2012), innate and adaptive immune responses (Ma *et al.*, 2011) and spinal motor neuron development (Otaegi *et al.*, 2011). Integrating miRNA-mediated RNA–RNA interactions is helpful for constructing more comprehensive regulatory networks. Our approach is flexible to analyze the integrated networks. With the increase in available miRNA expression profiles for AD and the accumulation of validated miRNA regulations, our results will become more precise. In addition, the approach was general and could be applied to other complex diseases.

ACKNOWLEDGEMENTS

We thank all the research staff for their contributions to this project and the anonymous reviewers for their helpful comments, which have greatly improved the manuscript.

Funding: Funds for Creative Research Groups of The National Natural Science Foundation of China [81121003]; the National Natural Science Foundation of China [30900837 and 81202074]; the Foundation for University Key Teacher of the Education Department of Heilongjiang Province [1252G037]; the Heilong Jiang Postdoctoral Funds for scientific research initiation [LBH-Q11042].

Conflict of Interest: none declared.

REFERENCES

Backes,C. *et al.* (2012) An integer linear programming approach for finding deregulated subgraphs in regulatory networks. *Nucleic Acids Res.*, **40**, e43.

Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.

Bartel,D.P. (2009) MicroRNAs: target recognition and regulatory functions. *Cell*, **136**, 215–233.

Barton,A. *et al.* (2004) Association of protein kinase C alpha (PRKCA) gene with multiple sclerosis in a UK population. *Brain*, **127**, 1717–1722.

Berezovska,O. *et al.* (1999) The Alzheimer-related gene presenilin 1 facilitates notch 1 in primary mammalian neurons. *Brain Res. Mol. Brain Res.*, **69**, 273–280.

Chen,C.D. *et al.* (2006) Visualization of APP dimerization and APP-Notch2 heterodimerization in living cells using bimolecular fluorescence complementation. *J. Neurochem.*, **97**, 30–43.

Chen,C.Y. *et al.* (2011) Coregulation of transcription factors and microRNAs in human transcriptional regulatory network. *BMC Bioinformatics*, **12(Suppl. 1)**, S41.

Chen,X.C. *et al.* (2003) Involvement of CDK4, pRB, and E2F1 in ginsenoside Rg1 protecting rat cortical neurons from beta-amyloid-induced apoptosis. *Acta Pharmacol. Sin.*, **24**, 1259–1264.

Cho,J.H. *et al.* (2011) An integrative approach to inferring biologically meaningful gene modules. *BMC Syst. Biol.*, **5**, 117.
Christensen,M.A. *et al.* (2004) Transcriptional regulation of BACE1, the beta-amyloid precursor protein beta-secretase, by Sp1. *Mol. Cell. Biol.*, **24**, 865–874.
Citron,B.A. *et al.* (2008) Transcription factor Sp1 dysregulation in Alzheimer's disease. *J. Neurosci. Res.*, **86**, 2499–2504.
Edgar,R. *et al.* (2002) Gene expression omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res.*, **30**, 207–210.
Ferrer,I. and Blanco,R. (2000) N-myc and c-myc expression in Alzheimer disease, Huntington disease and Parkinson disease. *Brain Res. Mol. Brain Res.*, **77**, 270–276.
Gaire,R.K. *et al.* (2013) Discovery and analysis of consistent active sub-networks in cancers. *BMC Bioinformatics*, **14**, S7.
Hebert,S.S. *et al.* (2008) Loss of microRNA cluster miR-29a/b-1 in sporadic Alzheimer's disease correlates with increased BACE1/beta-secretase expression. *Proc. Natl Acad. Sci. USA*, **105**, 6415–6420.
Holt,J. *et al.* (2009) Clinical inquiries. Do patients at high risk of Alzheimer's disease benefit from early treatment? *J. Fam. Pract.*, **58**, 320–322.
Hsu,S.D. *et al.* (2011) miRTarBase: a database curates experimentally validated microRNA-target interactions. *Nucleic Acids Res.*, **39**, D163–D169.
Ideker,T. *et al.* (2002) Discovering regulatory and signalling circuits in molecular interaction networks. *Bioinformatics*, **18(Suppl. 1)**, S233–S240.
Ikonen,M. *et al.* (2003) Interaction between the Alzheimer's survival peptide humanin and insulin-like growth factor-binding protein 3 regulates cell survival and apoptosis. *Proc. Natl Acad. Sci. USA*, **100**, 13042–13047.
Jaffrezic,F. *et al.* (2007) A structural mixed model for variances in differential gene expression studies. *Genet Res.*, **89**, 19–25.
Jiang,Q. *et al.* (2009) miR2Disease: a manually curated database for microRNA deregulation in human disease. *Nucleic Acids Res.*, **37**, D98–D104.
Junn,E. and Mouradian,M.M. (2012) MicroRNAs in neurodegenerative diseases and their therapeutic potential. *Pharmacol. Ther.*, **133**, 142–150.
Kantharidis,P. *et al.* (2011) Diabetes complications: the microRNA perspective. *Diabetes*, **60**, 1832–1837.
Keller,A. *et al.* (2009) A novel algorithm for detecting differentially regulated paths based on gene set enrichment analysis. *Bioinformatics*, **25**, 2787–2794.
Kitamura,Y. *et al.* (1997) Alteration of transcription factors NF-kappaB and STAT1 in Alzheimer's disease brains. *Neurosci. Lett.*, **237**, 17–20.
Krishnappa,R. (2011) Molecular expression profiling with respect to KEGG hsa05219 pathway. *Ecancermedicalscience*, **5**, 189.
Lahiri,D.K. *et al.* (2004) Apolipoprotein gene and its interaction with the environmentally driven risk factors: molecular, genetic and epidemiological studies of Alzheimer's disease. *Neurobiol. Aging*, **25**, 651–660.
Leininger-Muller,B. *et al.* (2003) Myeloperoxidase G-463A polymorphism and Alzheimer's disease in the ApoEurope study. *Neurosci. Lett.*, **349**, 95–98.
Li,X. *et al.* (2012) Dissection of human MiRNA regulatory influence to subpathway. *Brief. Bioinform.*, **13**, 175–186.
Liang,D. *et al.* (2012) Concerted perturbation observed in a hub network in Alzheimer's disease. *PLoS One*, **7**, e40498.
Liang,W.S. *et al.* (2007) Gene expression profiles in anatomically and functionally distinct regions of the normal aged human brain. *Physiol. Genomics*, **28**, 311–322.
Liu,Z.P. *et al.* (2010) Identifying dysfunctional crosstalk of pathways in various regions of Alzheimer's disease brains. *BMC Syst. Biol.*, **4(Suppl. 2)**, S11.
Lu,M. *et al.* (2008) An analysis of human microRNA and disease associations. *PLoS One*, **3**, e3420.
Lukiw,W.J. and Alexandrov,P.N. (2012) Regulation of complement factor H (CFH) by multiple miRNAs in Alzheimer's disease (AD) brain. *Mol. Neurobiol.*, **46**, 11–19.
Lukiw,W.J. *et al.* (2008) An NF-kappaB-sensitive micro RNA-146a-mediated inflammatory circuit in Alzheimer disease and in stressed human brain cells. *J. Biol. Chem.*, **283**, 31315–31322.
Ma,F. *et al.* (2011) The microRNA miR-29 controls innate and adaptive immune responses to intracellular bacterial infection by targeting interferon-gamma. *Nat. Immunol.*, **12**, 861–869.
Maloney,B. *et al.* (2010) Functional characterization of three single-nucleotide polymorphisms present in the human APOE promoter sequence: differential effects in neuronal cells and on DNA-protein interactions. *Am. J. Med. Genet. B Neuropsychiatr. Genet.*, **153B**, 185–201.
Marot,G. *et al.* (2009) Moderated effect size and P-value combinations for microarray meta-analyses. *Bioinformatics*, **25**, 2692–2699.
Meltzer,P.S. (2005) Cancer genomics: small RNAs with big impacts. *Nature*, **435**, 745–746.
Nitsch,D. *et al.* (2009) Network analysis of differential expression for the identification of disease-causing genes. *PLoS One*, **4**, e5526.
Nunez-Iglesias,J. *et al.* (2010) Joint genome-wide profiling of miRNA and mRNA expression in Alzheimer's disease cortex reveals altered miRNA regulation. *PLoS One*, **5**, e8898.
Otaegi,G. *et al.* (2011) An optimized sponge for microRNA miR-9 affects spinal motor neuron development *in vivo*. *Front. Neurosci.*, **5**, 146.
Piedrafita,F.J. *et al.* (1996) An Alu element in the myeloperoxidase promoter contains a composite SP1-thyroid hormone-retinoic acid response element. *J. Biol. Chem.*, **271**, 14412–14420.
Ramasamy,A. *et al.* (2008) Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med.*, **5**, e184.
Ramberg,V. *et al.* (2011) The CCAAT/enhancer binding protein (C/EBP) delta is differently regulated by fibrillar and oligomeric forms of the Alzheimer amyloid-beta peptide. *J. Neuroinflammation*, **8**, 34.
Ruan,J. *et al.* (2010) A general co-expression network-based approach to gene expression analysis: comparison and applications. *BMC Syst. Biol.*, **4**, 8.
Safran,M. *et al.* (2002) GeneCards 2002: towards a complete, object-oriented, human gene compendium. *Bioinformatics*, **18**, 1542–1543.
Santpere,G. *et al.* (2006) Abnormal Sp1 transcription factor expression in Alzheimer disease and tauopathies. *Neurosci. Lett.*, **397**, 30–34.
Sethupathy,P. *et al.* (2006) TarBase: A comprehensive database of experimentally supported animal microRNA targets. *RNA*, **12**, 192–197.
Shioya,M. *et al.* (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.
Sisodia,S.S. and St George-Hyslop,P.H. (2002) gamma-Secretase, Notch, Abeta and Alzheimer's disease: where do the presenilins fit in? *Nat. Rev. Neurosci.*, **3**, 281–290.
Sumazin,P. *et al.* (2011) An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. *Cell*, **147**, 370–381.
Thies,W. and Bleiler,L. (2011) 2011 Alzheimer's disease facts and figures. *Alzheimers Dement.*, **7**, 208–244.
Tuppo,E.E. and Arias,H.R. (2005) The role of inflammation in Alzheimer's disease. *Int. J. Biochem. Cell. Biol.*, **37**, 289–305.
Vaske,C.J. *et al.* (2010) Inference of patient-specific pathway activities from multi-dimensional cancer genomics data using PARADIGM. *Bioinformatics*, **26**, i237–i245.
Vlachos,I.S. *et al.* (2012) DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways. *Nucleic Acids Res.*, **40**, W498–W504.
Wang,J. *et al.* (2010) TransmiR: a transcription factor-microRNA regulation database. *Nucleic Acids Res.*, **38**, D119–D122.
Wang,W.X. *et al.* (2008) The expression of microRNA miR-107 decreases early in Alzheimer's disease and may accelerate disease progression through regulation of beta-site amyloid precursor protein-cleaving enzyme 1. *J. Neurosci.*, **28**, 1213–1223.
Wingender,E. *et al.* (2000) TRANSFAC: an integrated system for gene expression regulation. *Nucleic Acids Res.*, **28**, 316–319.
Xia,W. *et al.* (1997) Interaction between amyloid precursor protein and presenilins in mammalian cells: implications for the pathogenesis of Alzheimer disease. *Proc. Natl Acad. Sci. USA*, **94**, 8208–8213.
Xiao,F. *et al.* (2009) miRecords: an integrated resource for microRNA-target interactions. *Nucleic Acids Res.*, **37**, D105–D110.
Zhao,J. *et al.* (2011) Ranking candidate disease genes from gene expression and protein interaction: a Katz-centrality based approach. *PLoS One*, **6**, e24306.
Zhu,S. *et al.* (2012) The microRNA miR-23b suppresses IL-17-associated autoimmune inflammation by targeting TAB2, TAB3 and IKK-alpha. *Nat. Med.*, **18**, 1077–1086.