Research Article

Transcriptional Regulation of lncRNA Genes by Histone Modification in Alzheimer's Disease

Guoqiang Wan,[1] Wenyang Zhou,[1] Yang Hu,[1] Rui Ma,[1] Shuilin Jin,[2] Guiyou Liu,[1] and Qinghua Jiang[1]

[1]*School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001, China*
[2]*Department of Mathematics, Harbin Institute of Technology, Harbin 150001, China*

Correspondence should be addressed to Qinghua Jiang; qhjiang@hit.edu.cn

Increasing studies have revealed that long noncoding RNAs (lncRNAs) are not transcriptional noise but play important roles in the regulation of a wide range of biological processes, and the dysregulation of lncRNA genes is associated with disease development. Alzheimer's disease (AD) is a chronic neurodegenerative disease that usually starts slowly and gets worse over time. However, little is known about the roles of lncRNA genes in AD and how the lncRNA genes are transcriptionally regulated. Herein, we analyzed RNA-seq data and ChIP-seq histone modification data from CK-p25 AD model and control mice and identified 72 differentially expressed lncRNA genes, 4,917 differential peaks of H3K4me3, and 1,624 differential peaks of H3K27me3 between AD and control samples, respectively. Furthermore, we found 92 differential peaks of histone modification H3K4me3 are located in the promoter of 39 differentially expressed lncRNA genes and 8 differential peaks of histone modification H3K27me3 are located upstream of 7 differentially expressed lncRNA genes, which suggest that the majority of lncRNA genes may be transcriptionally regulated by histone modification in AD.

1. Introduction

Alzheimer's disease (AD) is a neurodegenerative disease with unknown etiology [1–3]. The main clinical manifestation is intelligence damage. In addition, it is the cause of 60% to 70% of cases of dementia. AD often begins in people over 65 years of age, and it affects approximately 6% of people aged 65 years and older [4]. There are about 48 million persons suffering from AD around the world in 2015, and dementia resulted in about 486,000 deaths in 2010 [5].

Long noncoding RNAs (lncRNAs) are non-protein-coding transcripts longer than 200 nucleotides in length. Thousands of human and mouse lncRNAs have been identified and emerging studies have revealed that lncRNAs play important roles in a wide range of biological processes and diseases [6–12]. Many studies have demonstrated that lncRNAs play crucial roles in the regulation of gene expression at epigenetic, transcriptional, and posttranscriptional level [13]. However, little is known about how lncRNA genes are transcriptionally regulated [14] in disease such as AD.

In this paper, we analyzed RNA-seq data and ChIP-seq histone modification data from control mice and CK-p25 AD model at 2 weeks after induction of neurodegeneration and checked whether lncRNA genes are transcriptionally regulated by histone modification in AD.

2. Materials and Methods

2.1. RNA-seq and ChIP-seq Data in AD and Control. The RNA-seq and ChIP-seq data were downloaded from GEO database with ID GSE65159 [15]. There are three control samples and three AD mice model samples at 2 weeks after induction of neurodegeneration. The histone modification marks include H3K4me3 and H3K27me3.

2.2. Identifying Differentially Expressed lncRNA Genes between AD and Control. We used RNA-seq data to evaluate gene expression on control mice and CK-p25 Alzheimer's disease model. We used the mm10 reference sequence to build an index by Bowtie2-build [16]; the mm10 reference sequence
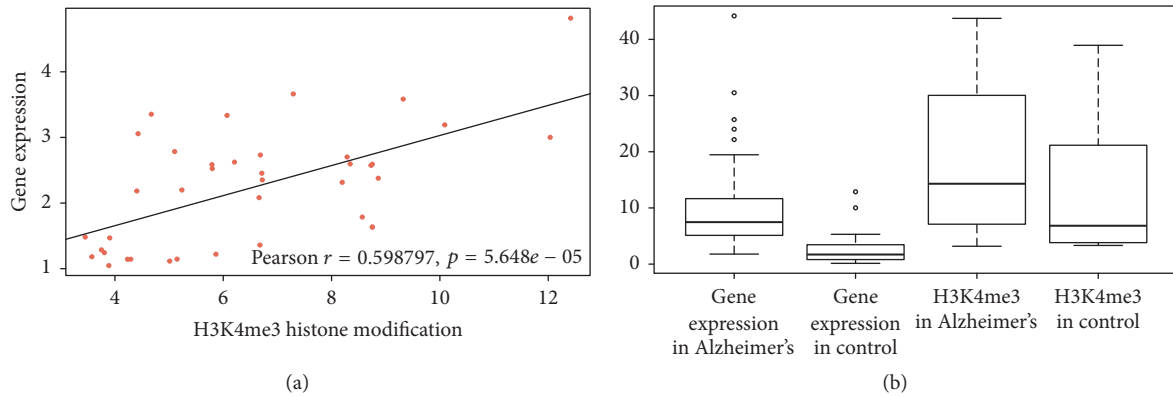
(a)

(b)

FIGURE 1: Positive association between expression level of differential lncRNA genes and H3K4me3 modification level in promoters of the differential lncRNA genes. (a) Scatter diagram and a fitting line show the positive association between fold change of lncRNA gene expression and fold change of H3K4me3 modification level in the promoter. (b) Boxplot of expression level of differential lncRNA genes and H3K4me3 modification level in the promoters of the differential lncRNA genes in AD and control samples, respectively, which shows that lncRNA genes with high H3K4me3 level in the promoters have high expression level. The circle in (b) refers to a singular point in statistics, differential from other points. But the singular point has statistical significance, showing the accuracy and objectivity of this article.

was downloaded from UCSC. Next the RNA-seq data are mapped to the mm10 reference sequence with TopHat2 [17] by default parameters. Cufflinks [18] was used to assemble the outcome of mapping and evaluate gene expression index. The lncRNA annotation was downloaded from GENCODE database, and differentially expressed lncRNA genes were identified by Cuffdiff with default parameters, a component of Cufflinks software.

2.3. Identifying Differential Histone Modification Peaks. To explore whether differentially expressed lncRNAs between AD and control are regulated by histone modification or not, we identified differential histone modification regions by analyzing the ChIP-seq data of histone marks H3K4me3 and H3K27me3 in AD and control. We firstly mapped the ChIP-seq data to the mm10 reference sequence by Bowtie2 software with default parameters. Then we used MACS2-callpeak [19] to identify the peaks of histone modification regions in the control mice and CKp25 Alzheimer's disease model [20], respectively. Finally, MACS2-bdgdiff is used to identify significantly differential histone modification regions between the control and AD.

2.4. Linking the Differential lncRNA Genes with the Differential Histone Modification Peaks Based on the Genomic Position. After identifying differential histone modification regions and differentially expressed lncRNA genes, we investigated whether the differential histone modification regions are located in the regulatory regions of the differential lncRNA genes. Herein, the regulatory regions are defined as 10 kbp upstream to 1 kbp downstream of transcriptional start site (TSS) of each differentially expressed lncRNA gene.

3. Results

3.1. Differentially Expressed lncRNA Genes between AD and Control Samples. By analyzing three AD and control RNA-seq data, we identified 72 significantly differentially expre-

ssed lncRNA genes with the BH-adjusted p value < 0.05 and fold change >2 (Supplementary Table 1, in Supplementary Material available online at http://dx.doi.org/10.1155/2016/3164238).

3.2. Differential Histone Modification Peaks between AD and Control Samples. We analyzed ChIP-seq histone modification data from CK-p25 AD model and control mice and identified 4,917 differential peaks of H3K4me3 and 1,624 differential peaks of H3K27me3 between AD and control samples, respectively.

3.3. Differential Histone Modification Peaks Are Located Upstream of Differentially Expressed lncRNA Genes. We found that there are 92 H3K4me3 differential histone modification peaks located in the promoters (2 kbp upstream to −1 kbp downstream) of 39 differentially expressed lncRNA genes (Supplementary Table 2) and 8 differential H3K27me3 histone modification peaks located in the region from 10 kb upstream to −1 kb downstream of 7 differentially expressed lncRNA genes. A positive association between histone modification level of H3K4me3 and lncRNA gene expression level is shown in Figure 1, and a negative association between histone modification level of H3K27me3 and lncRNA gene expression level is shown in Figure 2. A case study for the lncRNA gene named Gm20559 was shown in Figure 3, where the lncRNA Gm20559 had differential histone modification of H3K4me3 between AD and control in its promoter region, and exon 1 and exon 3 of Gm20559 are differentially expressed between AD and control. These results suggest that the majority of lncRNA genes (39 + 7)/72 may be transcriptionally regulated by histone modification in AD.

4. Discussion

lncRNA is a type of important regulatory RNAs that play critical roles in a wide range of biological processes. However, how the lncRNA genes themselves are transcriptionally regulated
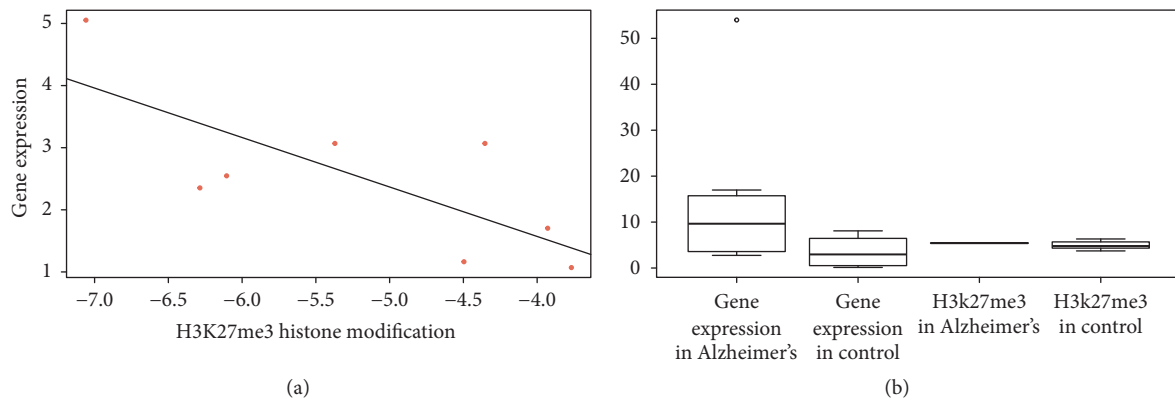
(a)

(b)

Figure 2: Negative association between expression level of differential lncRNA genes and H3K27me3 modification level in promoters of the differential lncRNA genes. (a) Scatter diagram and a fitting line show the negative association between fold change of expression level of differential lncRNA genes and fold change of H3K27me3 modification level in the promoters. (b) Boxplot of expression level of differential lncRNA genes and H3K27me3 modification level in the promoters of the differential lncRNA genes in AD and control samples, respectively, which shows that lncRNA genes with high H3K27me3 level in the promoters have low expression level. The circle in (b) refers to a singular point in statistics, differential from other points. But the singular point has statistical significance, showing the accuracy and objectivity of this article.
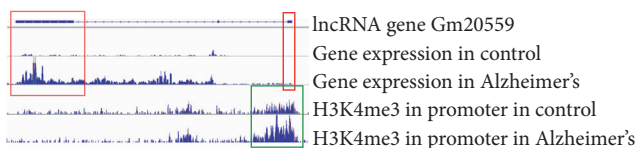

Figure 3: A lncRNA gene named Gm20559 with differential H3K4me3 modification level in the promoter between AD and control samples shows differential expression in exon 1 and exon 3. The red rectangle shows exon 1 and exon 3 regions of differentially expressed lncRNA gene Gm20559. And the green rectangle shows differential H3K4me3 histone modification in the promoter region, which suggests transcriptional regulation of Gm20559 by H3K4me3.

remains to be elucidated. In this paper, we used RNA-seq and ChIP-seq data from AD model and control to demonstrate that the majority of lncRNA genes are transcriptionally regulated by histone modification in AD.

As known, a protein-coding gene or lncRNA gene is regulated by many types of factors rather than one factor. Therefore, it sounds reasonable to integrate kinds of factors such as transcription factor, microRNA [21–24], DNA methylation, and histone modification to investigate the transcriptional regulation of lncRNAs in a specific condition such as AD, which will improve our understanding of lncRNA genes in AD.

Competing Interests

The authors declare no competing financial interests.

Authors' Contributions

Qinghua Jiang conceived the study. Guoqiang Wan collected the data, designed computational experiments, carried out statistical analysis, and wrote the manuscript. Wenyang Zhou

participated in the analysis of data. Yang Hu, Shuilin Jin, and Rui Ma participated in the revision of this manuscript. Guiyou Liu gave comments and revisions to the final version of this manuscript. All authors read and approved the final manuscript. Guoqiang Wan and Wenyang Zhou equally contributed to this paper.

Acknowledgments

References

[1] A. Burns and S. Iliffe, "Alzheimer's disease," *British Medical Journal*, vol. 338, article b158, 2009.

[2] G. Liu and Q. Jiang, "Alzheimer's disease *CD33* rs3865444 variant does not contribute to cognitive performance," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 113, no. 12, pp. E1589–E1590, 2016.

[3] G. Liu, Y. Xu, Y. Jiang, L. Zhang, R. Feng, and Q. Jiang, "PICALM rs3851179 variant confers susceptibility to alzheimer's disease in chinese population," *Molecular Neurobiology*, 2016.

[4] M. F. Mendez, "Early-onset Alzheimer's disease: nonamnestic subtypes and type 2 AD," *Archives of Medical Research*, vol. 43, no. 8, pp. 677–685, 2012.

[5] R. Lozano, M. Naghavi, K. Foreman et al., "Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010," *The Lancet*, vol. 380, no. 9859, pp. 2095–2128, 2012.

[6] X. Chen and G. Y. Yan, "Semi-supervised learning for potential human microRNA-disease associations inference," *Scientific Reports*, vol. 4, article 5501, 2014.

[7] X. Chen, C. C. Yan, X. Zhang, and Z.-H. You, "Long non-coding RNAs and complex diseases: from experimental results to computational models," *Briefings in Bioinformatics*, 2016.

[8] X. Chen and G.-Y. Yan, "Novel human lncRNA-disease association inference based on lncRNA expression profiles," *Bioinformatics*, vol. 29, no. 20, pp. 2617–2624, 2013.

[9] Q. Jiang, R. Ma, J. Wang et al., "LncRNA2Function: a comprehensive resource for functional investigation of human lncRNAs based on RNA-seq data," *BMC Genomics*, vol. 16, supplement 3, p. S2, 2015.

[10] Q. Jiang, J. Wang, X. Wu et al., "LncRNA2Target: a database for differentially expressed genes after lncRNA knockdown or overexpression," *Nucleic Acids Research*, vol. 43, no. 1, pp. D193–D196, 2015.

[11] G. Chen, Z. Wang, D. Wang et al., "LncRNADisease: a database for long-non-coding RNA-associated diseases," *Nucleic Acids Research*, vol. 41, no. 1, pp. D983–D986, 2013.

[12] X. Chen, M.-X. Liu, and G.-Y. Yan, "RWRMDA: predicting novel human microRNA-disease associations," *Molecular BioSystems*, vol. 8, no. 10, pp. 2792–2798, 2012.

[13] I. Ulitsky and D. P. Bartel, "LincRNAs: genomics, evolution, and mechanisms," *Cell*, vol. 154, no. 1, pp. 26–46, 2013.

[14] Q. Jiang, J. Wang, Y. Wang, R. Ma, X. Wu, and Y. Li, "TF2LncRNA: identifying common transcription factors for a list of lncRNA genes from ChIP-seq data," *BioMed Research International*, vol. 2014, Article ID 317642, 5 pages, 2014.

[15] E. Gjoneska, A. R. Pfenning, H. Mathys et al., "Conserved epigenomic signals in mice and humans reveal immune basis of Alzheimer's disease," *Nature*, vol. 518, no. 7539, pp. 365–369, 2015.

[16] B. Langmead and S. L. Salzberg, "Fast gapped-read alignment with Bowtie 2," *Nature Methods*, vol. 9, no. 4, pp. 357–359, 2012.

[17] S. Ghosh and C. K. Chan, "Analysis of RNA-seq data using TopHat and cufflinks," in *Plant Bioinformatics: Methods and Protocols*, vol. 1374 of *Methods in Molecular Biology*, pp. 339–361, Springer, Berlin, Germany, 2016.

[18] C. Trapnell, A. Roberts, L. Goff et al., "Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks," *Nature Protocols*, vol. 7, no. 3, pp. 562–578, 2012.

[19] Y. Zhang, T. Liu, C. A. Meyer et al., "Model-based analysis of ChIP-Seq (MACS)," *Genome Biology*, vol. 9, article R137, 2008.

[20] T. Liu, "Use model-based Analysis of ChIP-Seq (MACS) to analyze short reads generated by sequencing protein-DNA interactions in embryonic stem cells," in *Stem Cell Transcriptional Networks: Methods and Protocols*, B. L. Kidder, Ed., vol. 1150 of *Methods in Molecular Biology*, pp. 81–95, 2014.

[21] Q. Jiang, Y. Hao, G. Wang et al., "Prioritization of disease microRNAs through a human phenome-microRNAome network," *BMC Systems Biology*, vol. 4, supplement 1, article S2, 2010.

[22] Q. Jiang, G. Wang, S. Jin, Y. Li, and Y. Wang, "Predicting human microRNA-disease associations based on support vector machine," *International Journal of Data Mining and Bioinformatics*, vol. 8, no. 3, pp. 282–293, 2013.

[23] V. Agarwal, G. W. Bell, J.-W. Nam, and D. P. Bartel, "Predicting effective microRNA target sites in mammalian mRNAs," *eLife*, vol. 4, Article ID e05005, 2015.

[24] X. Chen, C. C. Yan, X. Zhang et al., "WBSMDA: within and between score for MiRNA-disease association prediction," *Scientific Reports*, vol. 6, article 21106, 2016.