ELSEVIER

Alzheimer's & Dementia ■ (2017) 1-14

Alzheimer's & Dementia

Featured Article

A systematic integrated analysis of brain expression profiles reveals *YAP1* and other prioritized hub genes as important upstream regulators in Alzheimer's disease

Min Xu[a,b,1], Deng-Feng Zhang[a,**,1], Rongcan Luo[a,b], Yong Wu[a,b], Hejiang Zhou[a], Li-Li Kong[a,c], Rui Bi[a], Yong-Gang Yao[a,b,d,*]

[a]*Key Laboratory of Animal Models and Human Disease Mechanisms of the Chinese Academy of Sciences & Yunnan Province, Kunming Institute of Zoology, Kunming, Yunnan, China*

[b]*Kunming College of Life Science, University of Chinese Academy of Sciences, Kunming, Yunnan, China*

[c]*Institute of Health Science, Anhui University, Hefei, Anhui, China*

[d]*CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China*

Abstract

Introduction: Profiling the spatial-temporal expression pattern and characterizing the regulatory networks of brain tissues are vital for understanding the pathophysiology of Alzheimer's disease (AD).

Methods: We performed a systematic integrated analysis of expression profiles of AD-affected brain tissues (684 AD and 562 controls). A network-based convergent functional genomic approach was used to prioritize possible regulator genes during AD development, followed by functional characterization.

Results: We generated a complete list of differentially expressed genes and hub genes of the transcriptomic network in AD brain and constructed a Web server (www.alzdata.org) for public access. Seventeen hub genes active at the early stages, especially *YAP1*, were recognized as upstream regulators of the AD network. Cellular assays proved that early alteration of *YAP1* could promote AD by influencing the whole transcriptional network.

Discussion: Early expression disturbance of hub genes is an important feature of AD development, and interfering with this process may reverse the disease progression.

Keywords: Alzheimer's disease; Differential expression; Convergent functional genomic; Upstream regulator; Database

1. Introduction

Alzheimer's disease (AD) is the most prevalent neurodegenerative disease in the elderly and is characterized by progressive memory loss and cognitive impairment. Pathological hallmarks of AD include the presence of extracellular amyloid-β (Aβ) plaques and intracellular neurofibrillary tangles, synaptic dysfunction, neuronal loss, and brain atrophy [1]. The occurrence and development of AD is affected by age, genetic, and environmental factors [2]. Previous linkage analyses have identified the Aβ production–related genes *APP*, *PSEN1*, and *PSEN2* as the causal genes for familial AD, whereas genome-wide association studies (GWASs) have identified two dozen of the susceptibility loci responsible for sporadic AD [3,4]. Despite the fact that remarkable advances have been made in the understanding of the genetic basis of AD, the pathophysiology of AD is not well understood. A complete characterization of the transcriptomic alterations and the regulatory mechanisms

The authors declare no conflict of interest.

[1]These authors contributed equally to this work.

*Corresponding author. Tel./Fax: +86-871-68125440.

**Corresponding author. Tel.: ■■■; Fax: ■■■.

E-mail address: zhangdengfeng@mail.kiz.ac.cn (D.-F.Z.), yaoyg@mail.kiz.ac.cn (Y.-G.Y.)

http://dx.doi.org/10.1016/j.jalz.2017.08.012

1552-5260/

that underpin AD may provide essential evidence to fill the gap.

Gene expression profiling of postmortem brain tissues from AD patients and normal controls has identified numerous dysregulated genes and contributed to the understanding of the biological processes disrupted during AD pathogenesis [5–31]. Among the list of differentially expressed genes (DEGs), dysfunction of mitochondrial pathways, calcium signaling, and neuroinflammation were consistently observed in AD, to name a few [32]. Nevertheless, the statistical power and consistency of previous individual studies were limited, mainly due to relatively small sample sizes [5–31]. A comprehensive, robust, cross-validated list of DEGs based on a large sample size is urgently needed in AD research. It is also important to understand the spatial-temporal expression pattern and regulatory network of these DEGs. Identifying the genes which play central regulatory roles during AD pathogenesis, and finding how these genes regulate the downstream DEGs, is vital for understanding both the pathophysiology of AD and looking for potential targets for drug therapy (Fig. 1A).

In this study, we performed an integrative analysis of available high-throughput brain expression profiling data sets from AD patients and controls using a convergent functional genomic (CFG) method, in an attempt to answer the aforementioned questions (Fig. 1B). We merged all the available expression data for four brain regions affected by AD (entorhinal cortex [EC], hippocampus [HP], temporal cortex [TC], and frontal cortex [FC]) [5–30] through cross-platform normalization to achieve the largest AD brain expression data set for these brain regions (1246 samples, including 139 HP, 78 EC, 697 FC, and 332 TC). We investigated the regulation pattern of DEGs in AD brain and prioritized hub genes and potential upstream regulator genes by the CFG method, which integrated various levels of AD-related data including GWAS, protein-protein interaction (PPI), brain expressional quantitative trait loci (eQTL), and expression data of mouse AD models. We have been able to provide a complete and robust list of DEGs and hub regulators and identified several candidate upstream regulators in DEG networks, such as *YAP1* in the glial cell differentiation module. Our functional experiments

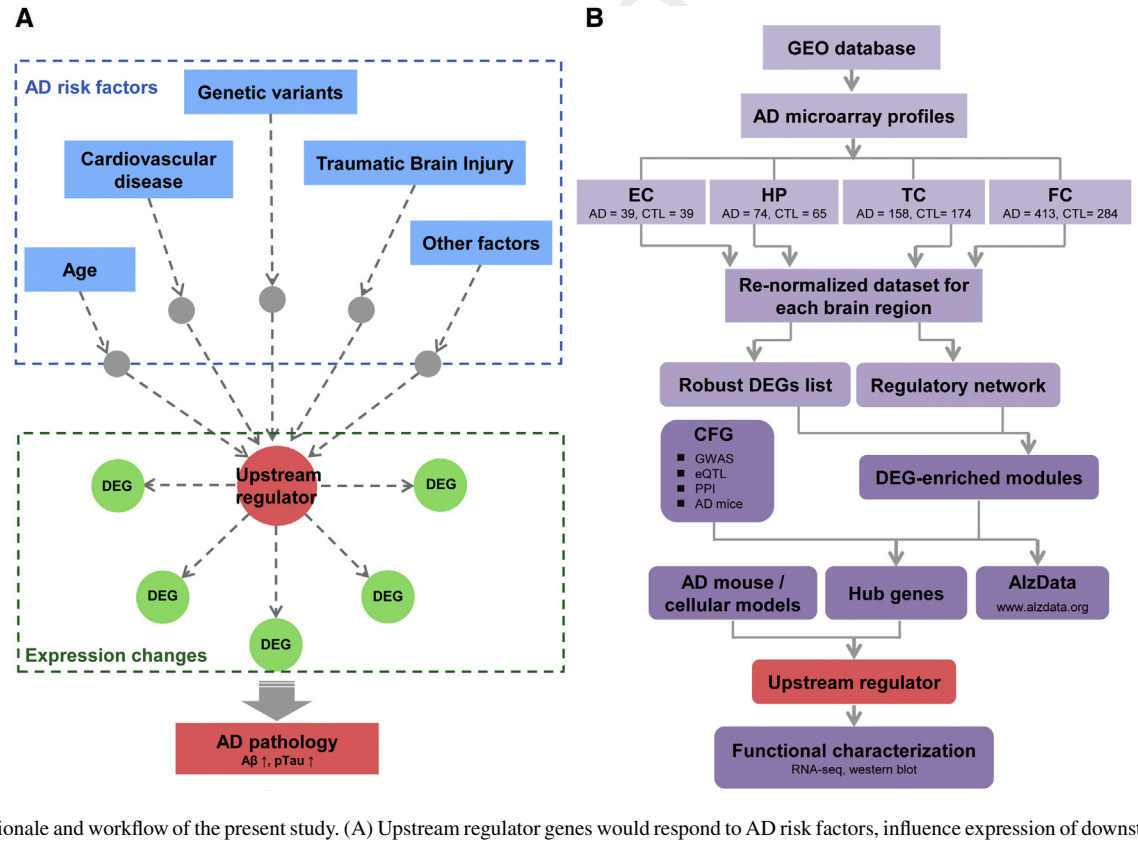

Fig. 1. Rationale and workflow of the present study. (A) Upstream regulator genes would respond to AD risk factors, influence expression of downstream genes, and thus promote AD pathology. Aβ ↑: increase of amyloid-beta level; pTau ↑: increase of phosphorylated tau. (B) We retrieved and renormalized all relevant expression data sets of samples with and without AD in entorhinal cortex (EC), hippocampus (HP), temporal cortex (TC), and frontal cortex (FC) and then analyzed for DEGs in the compiled data sets. We explored the gene prioritization and regulatory pattern in AD using coexpression network analysis and spatial-temporal expression data of AD mouse and cellular models and identified hub genes in the DEG-enriched coexpression networks. A convergent genomic approach (CFG) integrating multiple lines of evidence (including population genetic association, genetic regulation of expression, protein-protein interaction, early expression alteration, and pathology correlation in AD mice) was used to prioritize potential upstream genes. Abbreviations: eQTL, expressional quantitative trait loci; AD, Alzheimer's disease; DEG, differentially expressed gene; GEO, Gene Expression Omnibus; GWAS, genome-wide association study; PPI, protein-protein interaction.