Weighted Gene Co-expression Network Analysis Identifies Specific Modules and Hub Genes Related to Alzheimer's Disease

Kang Yang
Department of Neurology, Xiang'an Hospital of Xiamen University, Xiamen, Fujian Province
Bin Deng
Department of Anesthesiology, Xiang'an Hospital of Xiamen University, Xiamen, Fujian Province
Liju Xie
Department of Neurology, Xiang'an Hospital of Xiamen University, Xiamen, Fujian Province
Dandan Bo
Department of Neurology, Xiang'an Hospital of Xiamen University, Xiamen, Fujian Province
Jingsong Mao
Department of Radiology, Xiang'an Hospital of Xiamen University, Xiamen, Fujian Province
Yabin Song (ybsong@xah.xmu.edu.cn)
Department of Neurology, Xiang'an Hospital of Xiamen University, Xiamen, Fujian Province, China

Research Article

Keywords: Alzheimer's disease (AD), Weighted gene co-expression network (WGCNA), ATL1, CIRBP, MAPK

Posted Date: December 10th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-122428/v1

Abstract

Background: Alzheimer's disease (AD) is the most common cause of dementia. We used Weighted gene co-expression network (WGCNA) to construct a gene co-expression network to analyze AD-related genes and find biomarkers of AD. An AD RNA chip expression data was downloaded from NCBI's GEO database. The construction of the gene co-expression network was completed by the WGCNA package of R software. Functional annotation of genes was performed by gene ontology and Kyoto Encyclopedia Gene and Genomic pathway analysis.

Results: We used WGCNA to construct a co-expression network of 2713 differential expression genes in 56 AD patients and 44 healthy controls, and finally identified 7 gene modules, ranging in size from 32 to 1619. Correlation analysis of gene modules and disease showed that the turquoise module and the brown module were significantly associated with age. The turquoise module has obvious biological significance. The hub gene in this module is *ATL1*, which is involved in synaptic signaling and MAPK signaling pathways. The hub gene in the Brown module is *CIRBP*, which may be related to the regeneration of synapses.

Conclusions: This study used WGCNA, an efficient system biology algorithm, to analyze the hub genes related to AD for the first time and provide a new direction for the future research of AD.

Background

Alzheimer's disease (AD) is a degenerative disease of the nervous system characterized by clinical manifestations of memory impairment, aphasia, apraxia, agnosia, visual-spatial impairment, executive dysfunction, and changes in personality and behavior. It is estimated that in 2015, approximately 47.47 million people worldwide were diagnosed with dementia, and the number of people with dementia increased with the age of the population. These numbers are expected to increase from 50 million to 131 million by 2050[1].Although age is a major risk factor for AD, genetic factors account for approximately 60% to 80% of AD in the study of twins[2]. Alzheimer's disease is divided into familial AD and sporadic AD. *APOE4* is the main genetic risk factor for sporadic AD. *APOE4* homozygous carriers have a risk of AD exceeding 50%, and *APOE3* and *APOE4* heterozygous carriers have a risk of AD of 20%-30%. *APOE4* gene polymorphism indicates that the sporadic AD pathogenesis is complex. The most common familial AD is the mutation of *APP* on chromosome 21, *PSEN1* stained on chromosome 14, and *PSEN2* gene on chromosome 1. Although these signs are suspected to be closely related to the occurrence of AD, the exact pathogenesis of AD is still unclear[3].

The complex disease of AD is not caused only by a single genetic mutation, but may by the interaction of multiple minor genes. Common data analysis methods such as differential expression analysis performed only from a single genetic level can no longer meet the needs of researchers. From the perspective of systems biology, gene co-expression network analysis developed using graph theory technology provides an effective means to explore the interrelationship between biomolecules at the

system level. Weighted Gene Co-Expression Network Analysis (WGCNA) is a system biology algorithm for constructing a scale-free gene co-expression network that can be used to study the characteristics of gene networks in complex disease pathogenesis at the system level[4]. It not only categorizes different gene modules but also finds the relationship between clinical features and gene modules[5]. In this study, we used WGCNA to explore public GEO data and clinical information for AD patients. Several specific gene modules related to AD development were identified and several hub genes were discovered. These genes may play an important role in the pathogenesis or as a clinical biomarker for AD progression.

Methods

1. Total RNA chip data acquisition and screening differentially expressed genes （DEGs）

We apply the NCBI GEO database to download an RNA chip expression data GSE122063 for "Dementia Comparison: VaD vs. AD vs. Controls"(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE122063). GSE122063 is a qPCR whole-genome total RNA gene expression profiling was performed on frontal and temporal cortex from 56 AD patients, 36 Vascular dementia (VD) patients, and 44 healthy controls. Whole-genome gene expression was detected using an Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray 039381 (Feature Number version) chip.

All genes between AD and healthy controls were analyzed by t-test (p<0.05), and DEGs between the two were obtained for further construction of the co-expression network. The p value was corrected using the Bonferroni method.

2. WGCNA analysis of AD

2.1 Building a co-expression network

We constructed a weighted gene co-expression network analysis by the WGCNA package in the R platform[6]. First, we figured a pairwise similarity matrix for each set of genes and then used a soft threshold to transform the similarity matrix into an adjacency matrix. Secondly, to minimize the influence of noise and false correlation, we transformed the adjacency matrix into a topological overlap matrix and calculated the corresponding dissimilarity. Finally, a hierarchical clustering tree was drawn to present hierarchical clustering. In the tree diagram, each short vertical line corresponds to one gene, and each branch corresponds to one module. Since the branches of the tree group are densely interconnected, a hybrid dynamic tree cutting method was applied to cut the branches and divide the genes into different modules, and the thresholds for setting the minimum module size and the merge module were set to 34 and 0.15, respectively. Whether the gene can be selected to a module depends on the module membership assignment (kME). KME was determined by the correlation between the gene expression value and the module eigengene (ME), which represents the first principal component of the module. Assign genes with kME > 0.85 to modules with unique colors. Poor connectivity to other genes will be assigned to the gray module and will not be used for subsequent studies.

2.2 Association of Module and disease phenotype

The correlation between the module and the age and gender of AD patients was assessed by the Pearson correlation test. A module in which the ME exhibits a high positive or negative correlation with disease characteristics was selected as a candidate module to be studied.

2.3 Identify hub genes and biological annotations

Using the TopHub in R to calculate the hub gene in the candidate module, this step will calculate the most connected gene in each module with other genes. We regard that this gene plays a decisive role in the module and is called the hub gene[7]. Using metascape (http://metascape.org/gp/index.html) to perform gene ontology (GO) and Kyoto Encyclopedia Gene and Genomic (KEGG) pathway analysis on meaningful modular genes, setting the significance to *p*-value <0.01, a minimum value equal to 3, and enrichment factor > 1.5 (the enrichment factor is the count of observations and the count of accidental expectations).

Results

1. General statistics

A total of 100 patients were enrolled, including 56 AD and 44 healthy controls. The mean age of AD patients was 81 ± 6.6 years, and healthy control group was 79 ± 8.5 years. There was no significant difference in age between the two groups (*p*=0.166). There were 12 males (21.4%) and 44 females (78.6%) in the AD group, 20 males (45.5%) and 24 females (54.5%) in the control group.

2. Screening for DEG

The differential expression gene screening of 10,951 genes was performed by t-test, and the *p*-value was corrected by the Bonferroni method. Finally, 2,713 DEGs were selected.

3. Weighted gene co-expression network analysis

3.1 Building a network and identifying gene modules

For the network diagram of 2,713 genes, the weighting parameter β of the adjacency function is first selected according to the scale-free topology criterion in the WGCNA principle. The parameter β=8 was chosen to construct a co-expression network of the gene data set, and the correlation coefficient threshold was set to 0.85. The parameter selection diagram was drawn using R software (Figure 1). Next, using the step-by-step construction method in WGCNA, a hierarchical clustering tree is constructed to identify the co-expression network. Using the dynamic shear tree method to trim, set the gene cluster height to 0.85, and set deepSplit = 2 to ensure the uniqueness of the module and the balance with the surrounding gene. Modules with fewer genes are combined after initial module construction. As shown in Figure 2, a total of seven gene modules were identified, ranging in size from 32 to 1619(Table 1). Each

module was replaced with a different color, and gray indicates genes that could not be merged into any module.

3.2 Modules associated with disease traits and identification of hub genes

After determining the gene module, we further correlated the obtained modules with the clinical features of AD (age and gender). First, we calculate eigengene for each module, then associate eigengenes with external features and look for the most significant associations (Figure 3). Green indicates a module that is negatively correlated with disease, and red indicates a module that is positively correlated with disease. The results showed that the turquoise module (r=0.31, *p*=0.02) and the brown module (r=0.52, *p*=3E-05) were significantly positively correlated with the age of AD patients. Therefore, the brown module and the turquoise module will be used as candidate modules for further research. No obvious relevant modules were found in the gender aspect. The hub genes for the turquoise module and the brown module using TopHub calculations are *ATL1* and *CIRBP*, respectively.

4. Biological function annotation of genes in candidate module

Functional annotation and pathway analysis of all genes of two age-related candidate modules were performed by Metascape. The results suggest that the turquoise module is enriched with 18 GO terms and 2 KEGG pathways (Figure 4). The brown module is mainly enriched with 15 GO terms and 1 KEGG pathway (Figure 4). The turquoise module is primarily related to synapses, including synaptic signaling, regulation of postsynaptic membrane, synaptic transmission, glutamatergic, regulation of short-term neuronal synaptic plasticity. The KEGG pathway is MAPK signaling pathway. The brown module is mainly related to molecular metabolism and cell growth, such as the nucleotide metabolic process, mitochondrial transport, small molecule catabolic process.

Discussion

In this study, we used WGCNA method to construct a co-expression network of 2713 differential expression genes in 56 AD patients and 44 healthy controls, and finally identified 7 gene modules ranging in size from 32 to 1619. Correlation analysis of gene modules and disease phenotype showed that the turquoise module and the brown module were significantly associated with age. The turquoise module has obvious biological significance, which is related to synaptic signaling and MAPK signaling pathway. The hub gene in this module is *ATL1*. The Brown module is mainly related to molecular metabolism and cell growth. The hub gene of this module is *CIRBP*.

The algorithm of WGCNA software is to construct a gene co-expression network based on the similarity of expression profiles between samples, so as to systematically describe the gene expression information. It has advantages over traditional differential expression analysis and has been widely used in complex diseases. In this study, we identified seven gene modules using WGCNA, and the two modules with the strongest association with AD are the turquoise module and the brown module. The turquoise module has significant biological significance which is associated with synaptic signaling and MAPK

signaling pathways. Synapses are the basic unit of memory storage and information transmission in the brain. Previous analysis of brain samples from AD patients showed that a large number of synaptic losses are closely related to cognitive decline[8]. Various synaptic proteins, such as SNAP-25 (a presynaptic protein), PSD-95 (a post-synaptic protein), synapsin 1 and chromogranin B (synaptic vesicle protein), are reduced in AD patients brain[9, 10]. AD animal models also show defects in synaptic transmission as well as impairment of Long-term potentiation and Long-term depression[11]. These results indicate an imbalance in synaptic function in AD.

MAPK is a serine/threonine protein kinase. The MAPKs pathway is an important signal transduction pathway in mammalian cells. It is a key bridge connecting internal and external cells. It transduces extracellular stimulation signals into cells and their nucleus and causes cellular biological response. It is involved in physiological processes such as cell proliferation, differentiation, apoptosis and stress response. In mammals, several different MAPKs have been identified, including p38 MAPK, *c-junN*-terminal kinase (JNK), extracellular signal-regulated kinase (EERK1/2) and ERK5/BMK-1. Amongst MAPK, p38 MAPK is widely involved in signaling pathways of different biological functions. In the central nervous system, p38 MAPK is highly expressed in areas critical for learning and memory and may be a key component of advanced brain function[12]. By studying the brain tissue of the control group and AD patients after death, it was found that p38 MAPK was activated in the early stage of AD[13, 14]. Besides, the upstream activation molecule MAPK6 of p38 MAPK was also found to be up-regulated in the brain tissue of autopsy in AD patients[15]. Several studies have found that pericytes are the main component of cerebral vascular composition, and they are degraded and reduced in the hippocampus and cortex of AD patients[16-18]. Xu et al. found that Aβ inhibits the differentiation of mesenchymal stem cells into pericytes by activating ERK1/2MAPK signaling pathway[19]. Recent studies have confirmed that autophagy plays a role in AD[20]. Interestingly, MAPK has a special function as a positive regulator and a negative regulator of autophagy[21, 22]. MAPK can promote autophagy through phosphorylation of its BCL2[23]. These findings suggest that MAPK may be involved in the pathogenesis of AD. This is consistent with our results and reflects the reliability of WGCNA for identifying gene modules.

We use the TopHub in R to calculate the hub gene in the candidate module. The hub gene of the turquoise module is *ATL1*. Atlastin (ATL) belongs to the dynamin GTPase superfamily and is present in all vertebrates, as well as homologous proteins in many other organisms[24]. Novel studies have found that membrane protein atlastin is very important for homologous membrane fusion of tubular endoplasmic reticulum[25, 26]. In neurons, the endoplasmic reticulum is tubular in the axons[27], and in the dendrites, it contains a network and a tubular form[28, 29]. Impaired endoplasmic reticulum function is associated with many diseases, such as hereditary spastic paraplegia(HSP). HSP is a degenerative disease of the nervous system. It is classified into the simple type and variant type according to clinical features. The simple type is characterized by bilateral lower extremity spastic paralysis. Variants may include other nervous system manifestations such as optic atrophy, retinal pigmentation, epilepsy, Deafness and neurodevelopmental delay[30]. More than 50 genetic loci are genetics of HSP[31]. Interestingly, many mutations in *ATL-1* were found in HSP patients, suggesting that neuronal ER morphology may be associated with HSP pathogenesis[32]. However, neuronal ER injury, the relationship

between ATL-1 and HSP has not been determined, and how ATL-1 causes a neurodevelopmental delay in patients remains to be further studied. Interestingly, some studies have found that genetic mutations encoding spastin-SPAST can cause not only HSP but also AD. In HSP with SPAST mutation, spastin cuts down microtubule activity[33], whereas, in AD patient cell model, spastin is abnormally activated and induces microtubule breakdown[34]. Therefore, whether there is similar pathogenesis between HSP and AD, whether *ATL1* is also associated with the pathogenesis of AD remains to be further studied, which may provide a new direction for the diagnosis and treatment of AD.

CIRBP (cold-inducible RNA-binding protein) is an RNA-binding protein, also known as nuclear heterogeneous ribonucleoprotein, which is an 18kD protein of the glycine-rich RNA-binding protein family[35]. *CIRBP* is the first cold shock protein found in mammalian cells, and it can regulate cell growth and apoptosis under cold induction[36]. Similar to *CIRBP*, *RBM3* (RNA-binding morif protein 3), which is also a cold-induced protein, is up-regulated under cold stimulation and can regulate transcription and translation[37] . The reduction in the number of synapses is an early feature of neurodegenerative diseases. Diego Peretti[38] found that in prion-infected mice and 5XFAD (AD) mice, the ability to regenerate synapses after hypothermia is reduced, which is related to the failure of *RBM3* induction. *RBM3* over-expression can be achieved by reducing the temperature to increase endogenous levels or by lentiviral delivery before *RBM3* loses response, which can provide continuous synaptic protection for 5XFAD mice and prion-infected mice, thereby preventing behavioral defects and neurons lost and significantly prolonged survival. Therefore, the cold shock pathway can be used as a potential protective therapy in neurodegenerative diseases. Although much is known about the processes leading to synaptic dysfunction and loss, how it affects synaptic regeneration is still unknown. *CIRBP*, which is also a cold-inducible protein, can protect neuronal cells when its expression is increased. Whether it can also provide protection to synapses like *RBM3* is yet to be further studied. It may bring new therapeutic targets for the neuroprotection of neurodegenerative diseases.

Our results indicate that the identification of specific modules and hub genes in AD by WGCNA provides new clues for further research in the future. But our study also has some limitations. Firstly, AD can be divided into early-onset and late-onset according to the age of onset. The pathogenesis of these two types is not consistent. As we lacked more clinical data, we did not classify AD in our study. Secondly, we identified the key genes related to AD from microarray data analysis, which should be further verified in vitro or vivo experiments. Although we found some positive results, the sample size is still small. A large number of clinical samples will be needed in the future to validate our clinical results and to clarify the underlying mechanisms by which key genes affect AD.

Conclusions

In short, the development of AD is a complex pathological process in which multiple factors work together. It is necessary to use WGCNA to identify AD risk factors or biomarkers for therapeutic targets. Detection of specific gene modules will help elucidate the mechanisms of disease and promote the treatment of AD. The two important gene modules in the AD group involve a variety of biological

processes that can provide candidate targets for diagnosis and treatment. Besides, this study reveals that the differential expression of *ATL-1* gene may be related to the pathogenesis of AD. The cold shock protein *CIRBP* may help repair of synapses, providing new research directions for future research and provide valuable information for the treatment of AD.

Abbreviations

AD: Alzheimer's disease; WGCNA: Weighted gene co-expression network; GO: Gene Ontology; KEGG: Kyoto Encyclopedia Gene and Genomic; DEG: differentially expressed gene; VD: Vascular dementia; ME: Module eigengene; JNK: *c-junN*-terminal kinase; ATL: Atlastin; HSP: hereditary spastic paraplegia; *CIRBP*: cold-inducible RNA-binding protein; *RBM3*: RNA-binding morif protein 3.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Competing interests

The authors declare no conflict of interests.

Funding

This study was supported by the National Nature Science Foundation of China (81501207).

Authors' contributions

Kang Yang: study concept and design, statistical analysis, manuscript writing. Bin Deng: analysis and interpretation of data, critical revisions. Liju Xie, Dandan Bo and Jingsong Mao: analysis and interpretation of data. Yabin Song: study concept and design, manuscript writing.

Acknowledgements

The authors sincerely thank Erin Catherine Mckay's team for the data.

References

1. Winblad B, Amouyel P, Andrieu S, Ballard C, Brayne C, Brodaty H, Cedazo-Minguez A, Dubois B, Edvardsson D, Feldman H *et al*: **Defeating Alzheimer's disease and other dementias: a priority for European science and society**. *The Lancet Neurology* 2016, **15**(5):455-532.
2. Gatz M, Reynolds CA, Fratiglioni L, Johansson B, Mortimer JA, Berg S, Fiske A, Pedersen NL: **Role of genes and environments for explaining Alzheimer disease**. *Archives of general psychiatry* 2006, **63**(2):168-174.
3. Huang Y, Mucke L: **Alzheimer mechanisms and therapeutic strategies**. *Cell* 2012, **148**(6):1204-1222.
4. Presson AP, Sobel EM, Papp JC, Suarez CJ, Whistler T, Rajeevan MS, Vernon SD, Horvath S: **Integrated weighted gene co-expression network analysis with an application to chronic fatigue syndrome**. *BMC systems biology* 2008, **2**:95.
5. Langfelder P, Horvath S: **WGCNA: an R package for weighted correlation network analysis**. *BMC bioinformatics* 2008, **9**:559.
6. Zhang B, Horvath S: **A general framework for weighted gene co-expression network analysis**. *Stat Appl Genet Mol Biol* 2005, **4**:Article17.
7. Gaiteri C, Ding Y, French B, Tseng GC, Sibille E: **Beyond modules and hubs: the potential of gene coexpression networks for investigating molecular mechanisms of complex brain disorders**. *Genes Brain Behav* 2014, **13**(1):13-24.
8. Li K, Wei Q, Liu FF, Hu F, Xie AJ, Zhu LQ, Liu D: **Synaptic Dysfunction in Alzheimer's Disease: Abeta, Tau, and Epigenetic Alterations**. *Molecular neurobiology* 2018, **55**(4):3021-3032.
9. Marksteiner J, Kaufmann WA, Gurka P, Humpel C: **Synaptic proteins in Alzheimer's disease**. *Journal of molecular neuroscience : MN* 2002, **18**(1-2):53-63.
10. Masliah E, Mallory M, Alford M, DeTeresa R, Hansen LA, McKeel DW, Jr., Morris JC: **Altered expression of synaptic proteins occurs early during progression of Alzheimer's disease**. *Neurology* 2001, **56**(1):127-129.
11. Mota SI, Ferreira IL, Rego AC: **Dysfunctional synapse in Alzheimer's disease - A focus on NMDA receptors**. *Neuropharmacology* 2014, **76 Pt A**:16-26.
12. Yasuda S, Sugiura H, Tanaka H, Takigami S, Yamagata K: **p38 MAP kinase inhibitors as potential therapeutic drugs for neural diseases**. *Central nervous system agents in medicinal chemistry* 2011, **11**(1):45-59.
13. Pei JJ, Braak E, Braak H, Grundke-Iqbal I, Iqbal K, Winblad B, Cowburn RF: **Localization of active forms of C-jun kinase (JNK) and p38 kinase in Alzheimer's disease brains at different stages of neurofibrillary degeneration**. *Journal of Alzheimer's disease : JAD* 2001, **3**(1):41-48.
14. Sun A, Liu M, Nguyen XV, Bing G: **P38 MAP kinase is activated at early stages in Alzheimer's disease brain**. *Experimental neurology* 2003, **183**(2):394-405.
15. Zhu X, Rottkamp CA, Hartzler A, Sun Z, Takeda A, Boux H, Shimohama S, Perry G, Smith MA: **Activation of MKK6, an upstream activator of p38, in Alzheimer's disease**. *Journal of neurochemistry*

2001, **79**(2):311-318.

16. Halliday MR, Rege SV, Ma Q, Zhao Z, Miller CA, Winkler EA, Zlokovic BV: **Accelerated pericyte degeneration and blood-brain barrier breakdown in apolipoprotein E4 carriers with Alzheimer's disease**. *Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism* 2016, **36**(1):216-227.
17. Janota CS, Brites D, Lemere CA, Brito MA: **Glio-vascular changes during ageing in wild-type and Alzheimer's disease-like APP/PS1 mice**. *Brain research* 2015, **1620**:153-168.
18. Sagare AP, Bell RD, Zhao Z, Ma Q, Winkler EA, Ramanathan A, Zlokovic BV: **Pericyte loss influences Alzheimer-like neurodegeneration in mice**. *Nature communications* 2013, **4**:2932.
19. Xu L, Li J, Luo Z, Wu Q, Fan W, Yao X, Li Q, Yan H, Wang J: **Abeta inhibits mesenchymal stem cell-pericyte transition through MAPK pathway**. *Acta biochimica et biophysica Sinica* 2018, **50**(8):776-781.
20. Lee S, Sato Y, Nixon RA: **Primary lysosomal dysfunction causes cargo-specific deficits of axonal transport leading to Alzheimer-like neuritic dystrophy**. *Autophagy* 2011, **7**(12):1562-1563.
21. Xu P, Das M, Reilly J, Davis RJ: **JNK regulates FoxO-dependent autophagy in neurons**. *Genes & development* 2011, **25**(4):310-322.
22. Waetzig V, Herdegen T: **Context-specific inhibition of JNKs: overcoming the dilemma of protection and damage**. *Trends in pharmacological sciences* 2005, **26**(9):455-461.
23. Wei Y, Pattingre S, Sinha S, Bassik M, Levine B: **JNK1-mediated phosphorylation of Bcl-2 regulates starvation-induced autophagy**. *Molecular cell* 2008, **30**(6):678-688.
24. Zhao X, Alvarado D, Rainier S, Lemons R, Hedera P, Weber CH, Tukel T, Apak M, Heiman-Patterson T, Ming L *et al*: **Mutations in a newly identified GTPase gene cause autosomal dominant hereditary spastic paraplegia**. *Nature genetics* 2001, **29**(3):326-331.
25. Hu J, Shibata Y, Zhu PP, Voss C, Rismanchi N, Prinz WA, Rapoport TA, Blackstone C: **A class of dynamin-like GTPases involved in the generation of the tubular ER network**. *Cell* 2009, **138**(3):549-561.
26. Orso G, Pendin D, Liu S, Tosetto J, Moss TJ, Faust JE, Micaroni M, Egorova A, Martinuzzi A, McNew JA *et al*: **Homotypic fusion of ER membranes requires the dynamin-like GTPase atlastin**. *Nature* 2009, **460**(7258):978-983.
27. Yalcin B, Zhao L, Stofanko M, O'Sullivan NC, Kang ZH, Roost A, Thomas MR, Zaessinger S, Blard O, Patto AL *et al*: **Modeling of axonal endoplasmic reticulum network by spastic paraplegia proteins**. *eLife* 2017, **6**.
28. Cooney JR, Hurlburt JL, Selig DK, Harris KM, Fiala JC: **Endosomal compartments serve multiple hippocampal dendritic spines from a widespread rather than a local store of recycling membrane**. *The Journal of neuroscience : the official journal of the Society for Neuroscience* 2002, **22**(6):2215-2224.
29. Spacek J, Harris KM: **Three-dimensional organization of smooth endoplasmic reticulum in hippocampal CA1 dendrites and dendritic spines of the immature and mature rat**. *The Journal of*

neuroscience : the official journal of the Society for Neuroscience 1997, **17**(1):190-203.

30. Shribman S, Reid E, Crosby AH, Houlden H, Warner TT: **Hereditary spastic paraplegia: from diagnosis to emerging therapeutic approaches**. *The Lancet Neurology* 2019.
31. Noreau A, Dion PA, Rouleau GA: **Molecular aspects of hereditary spastic paraplegia**. *Experimental cell research* 2014, **325**(1):18-26.
32. Guelly C, Zhu PP, Leonardis L, Papic L, Zidar J, Schabhuttl M, Strohmaier H, Weis J, Strom TM, Baets J *et al*: **Targeted high-throughput sequencing identifies mutations in atlastin-1 as a cause of hereditary sensory neuropathy type I**. *American journal of human genetics* 2011, **88**(1):99-105.
33. Shoukier M, Neesen J, Sauter SM, Argyriou L, Doerwald N, Pantakani DV, Mannan AU: **Expansion of mutation spectrum, determination of mutation cluster regions and predictive structural classification of SPAST mutations in hereditary spastic paraplegia**. *European journal of human genetics : EJHG* 2009, **17**(2):187-194.
34. Zempel H, Luedtke J, Kumar Y, Biernat J, Dawson H, Mandelkow E, Mandelkow EM: **Amyloid-beta oligomers induce synaptic damage via Tau-dependent microtubule severing by TTLL6 and spastin**. *The EMBO journal* 2013, **32**(22):2920-2937.
35. Nishiyama H, Itoh K, Kaneko Y, Kishishita M, Yoshida O, Fujita J: **A glycine-rich RNA-binding protein mediating cold-inducible suppression of mammalian cell growth**. *The Journal of cell biology* 1997, **137**(4):899-908.
36. Liao Y, Tong L, Tang L, Wu S: **The role of cold-inducible RNA binding protein in cell stress response**. *International journal of cancer* 2017, **141**(11):2164-2173.
37. Lleonart ME: **A new generation of proto-oncogenes: cold-inducible RNA binding proteins**. *Biochimica et biophysica acta* 2010, **1805**(1):43-52.
38. Peretti D, Bastide A, Radford H, Verity N, Molloy C, Martin MG, Moreno JA, Steinert JR, Smith T, Dinsdale D *et al*: **RBM3 mediates structural plasticity and protective effects of cooling in neurodegeneration**. *Nature* 2015, **518**(7538):236-239.

Tables

Table 1. The number of genes in 8 modules

Module	Gene number
Blue	633
Brown	232
Green	67
Grey	32
Red	48
Turquoise	1619
Yellow	82

Figures

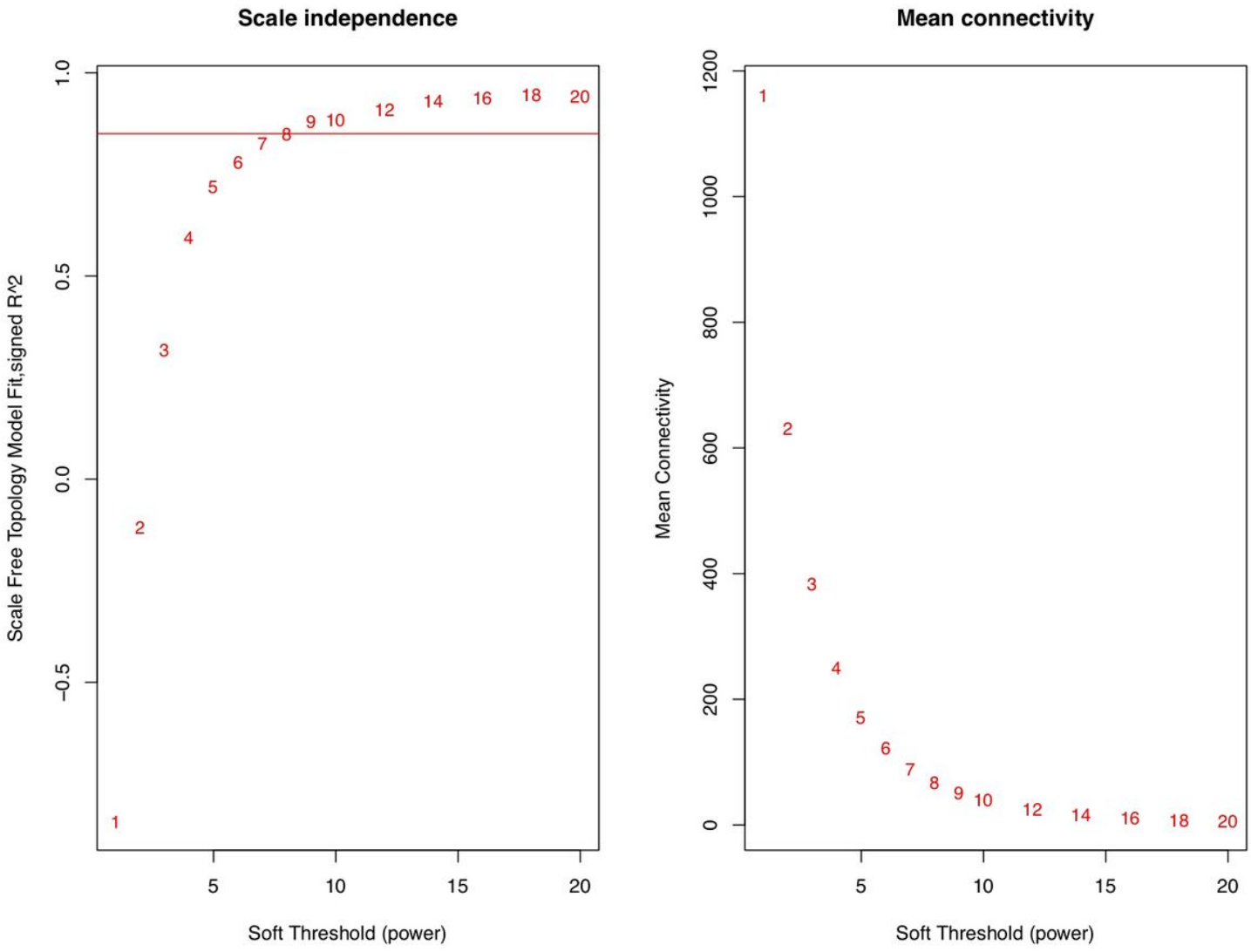

Figure 1

WGCNA threshold selection. The left graph shows the relationship between the soft threshold β and the correlation coefficient. The level of the correlation coefficient represents whether it conforms to the scale-free topology network structure model. The x-axis represents the soft threshold parameter and the y-axis represents the scale-free fit index. The right graph represents the correlation between the soft threshold β and the number of gene connections, indicating the average connection level of the gene network. The x-axis represents the soft threshold parameter and the y-axis represents the average connectivity.

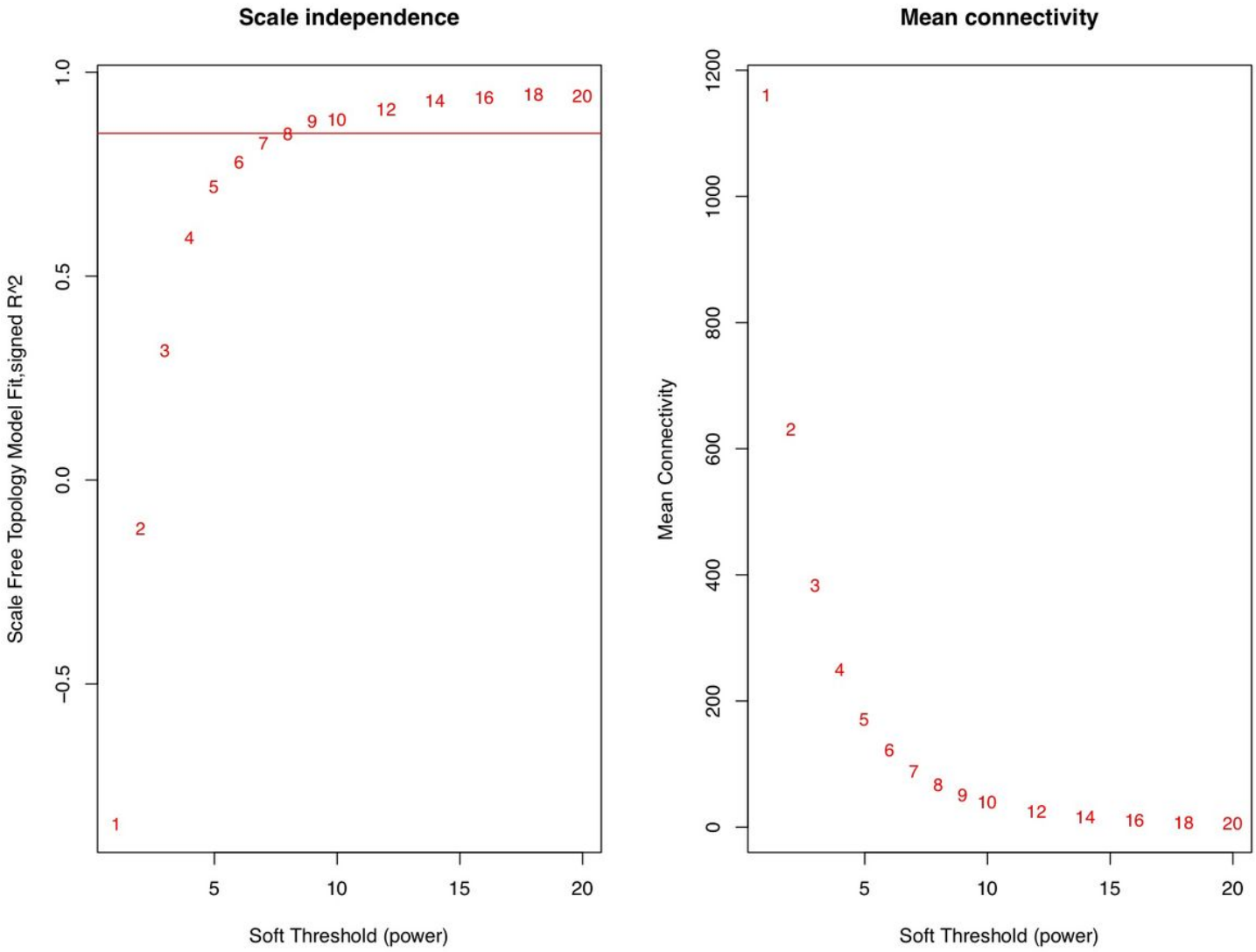

Figure 1

WGCNA threshold selection. The left graph shows the relationship between the soft threshold β and the correlation coefficient. The level of the correlation coefficient represents whether it conforms to the scale-free topology network structure model. The x-axis represents the soft threshold parameter and the y-axis represents the scale-free fit index. The right graph represents the correlation between the soft threshold β and the number of gene connections, indicating the average connection level of the gene network. The x-axis represents the soft threshold parameter and the y-axis represents the average connectivity.

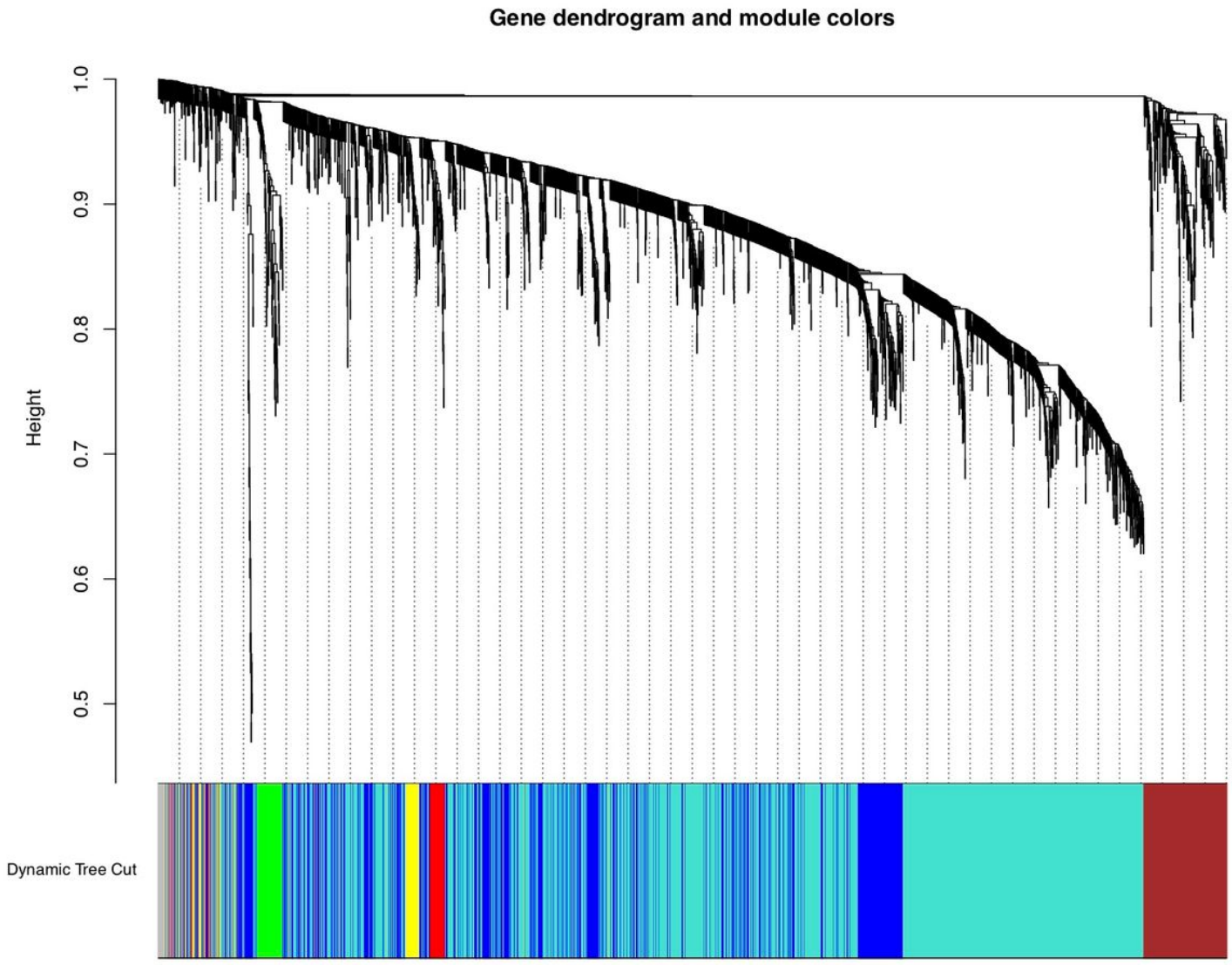

Figure 2

Clustering dendrograms of genes. A gene clustering tree (dendritic map) obtained by hierarchical clustering based on adjacent dissimilarity. The colored lines below the tree represent the module memberships identified by the dynamic tree cutting method. Height indicates the co-expression distance. Each color represents a module, and each vertical line represents a gene. A total of 7 gene modules were identified.

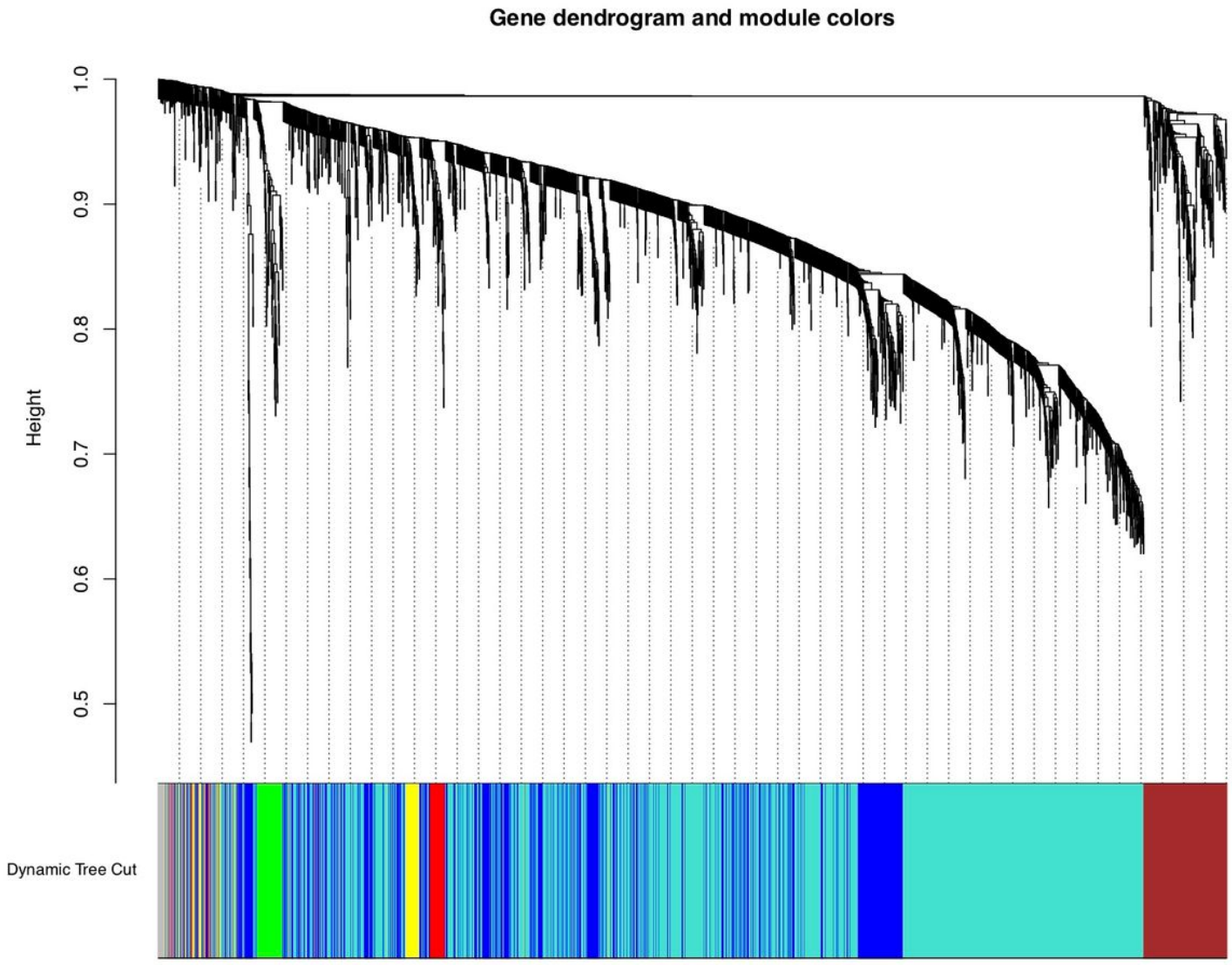

Figure 2

Clustering dendrograms of genes. A gene clustering tree (dendritic map) obtained by hierarchical clustering based on adjacent dissimilarity. The colored lines below the tree represent the module memberships identified by the dynamic tree cutting method. Height indicates the co-expression distance. Each color represents a module, and each vertical line represents a gene. A total of 7 gene modules were identified.

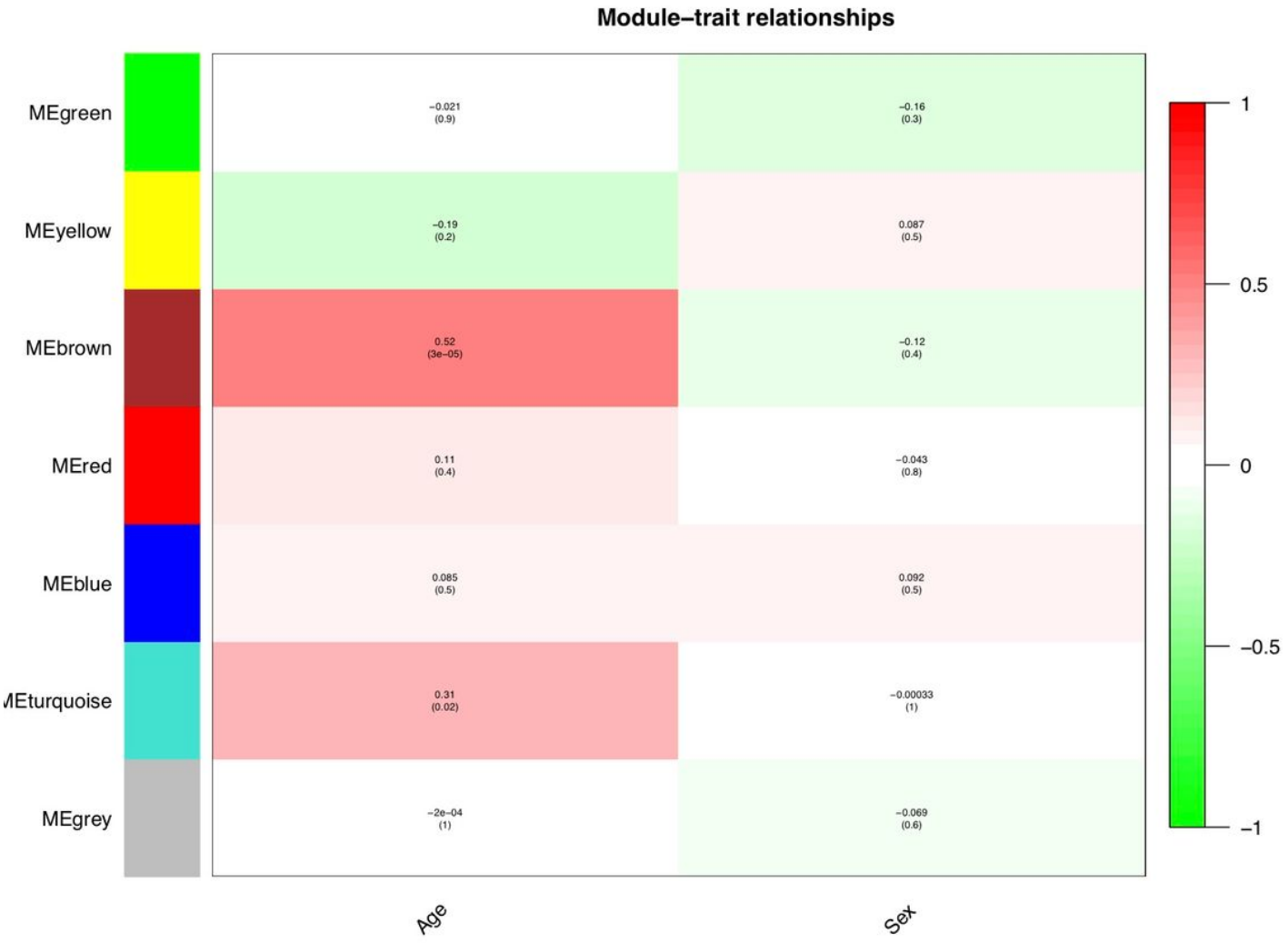

Figure 3

Module-disease association map. Each row corresponds to a module eigengene and each column corresponds to age and gender. Each cell contains the corresponding correlation and p-value.

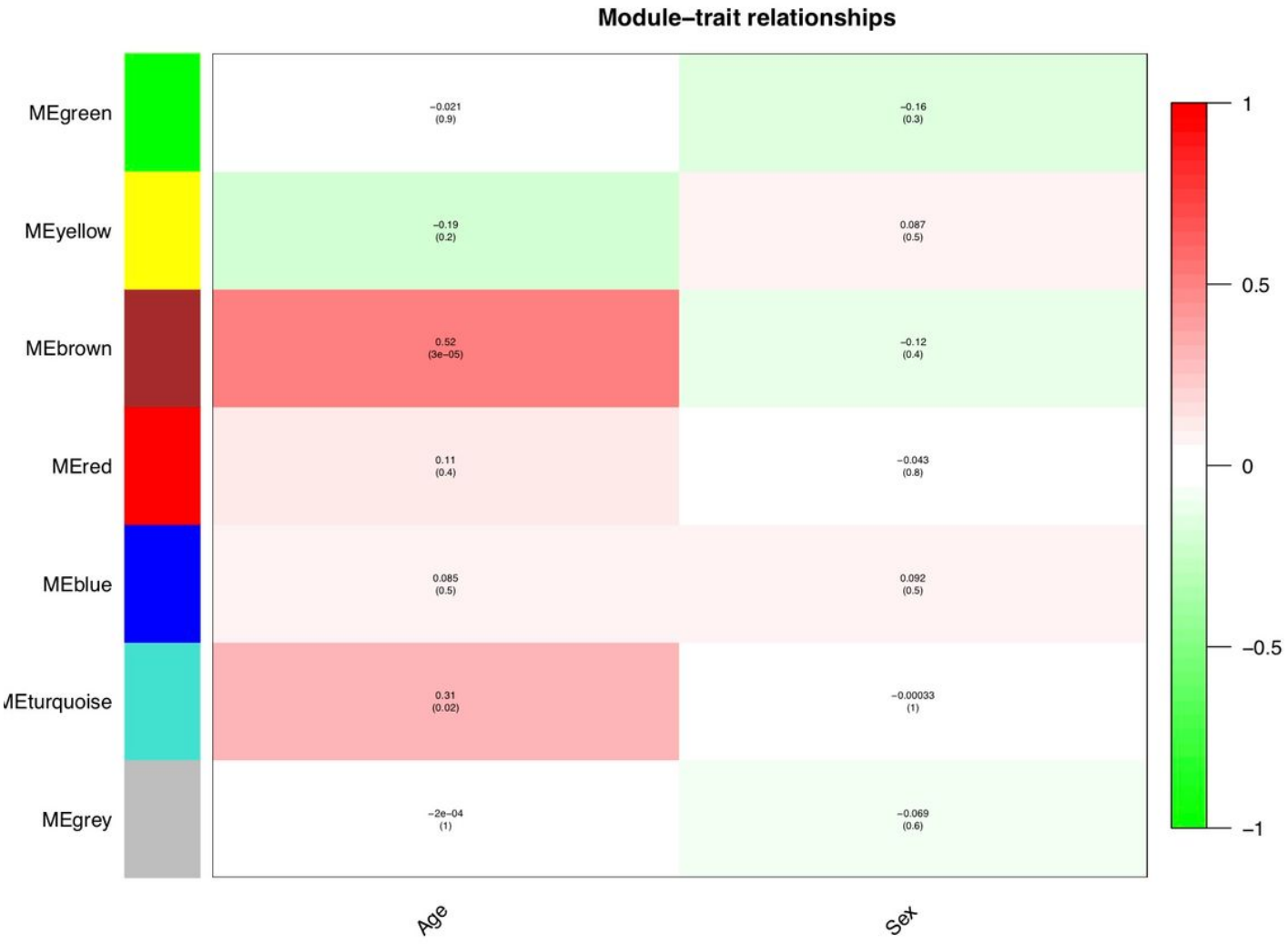

Figure 3

Module-disease association map. Each row corresponds to a module eigengene and each column corresponds to age and gender. Each cell contains the corresponding correlation and p-value.

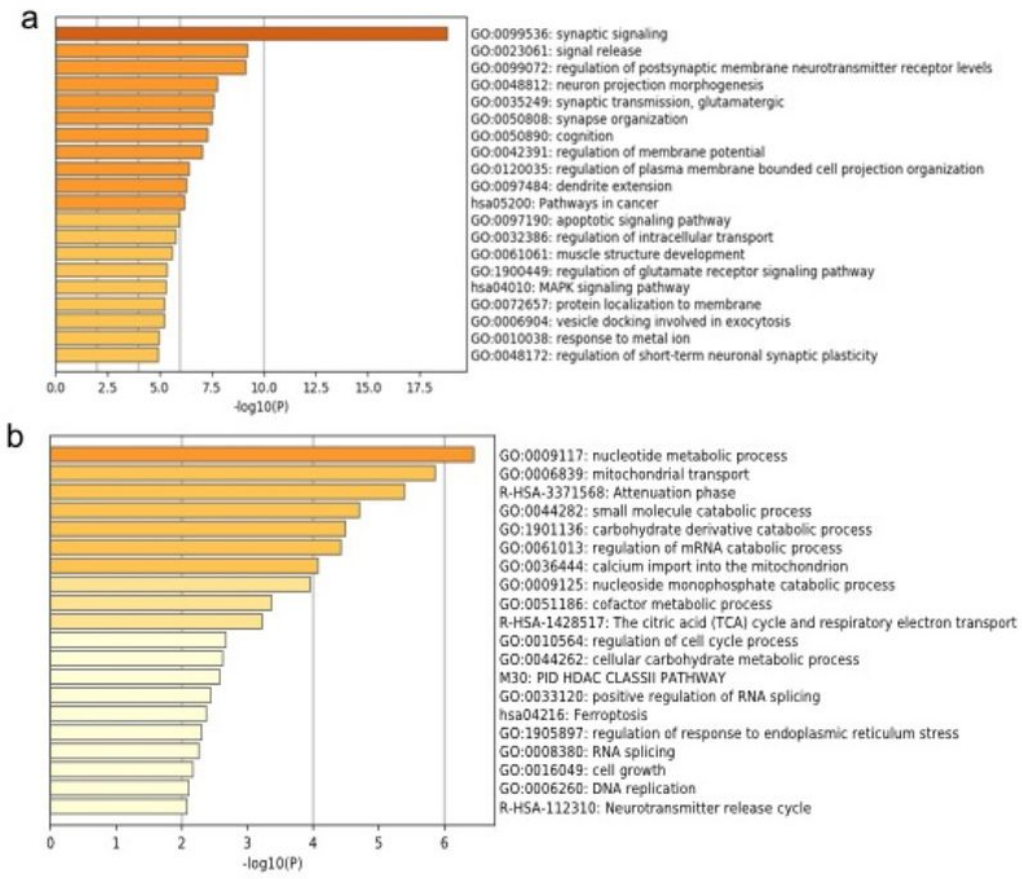

Figure 4

Plot of the enriched GO and KEGG terms in key co-expression modules. a: The plot of enriched biological process and KEEG pathway in the turquoise module. b: The plot of enriched biological process and KEEG pathway in the brown module.

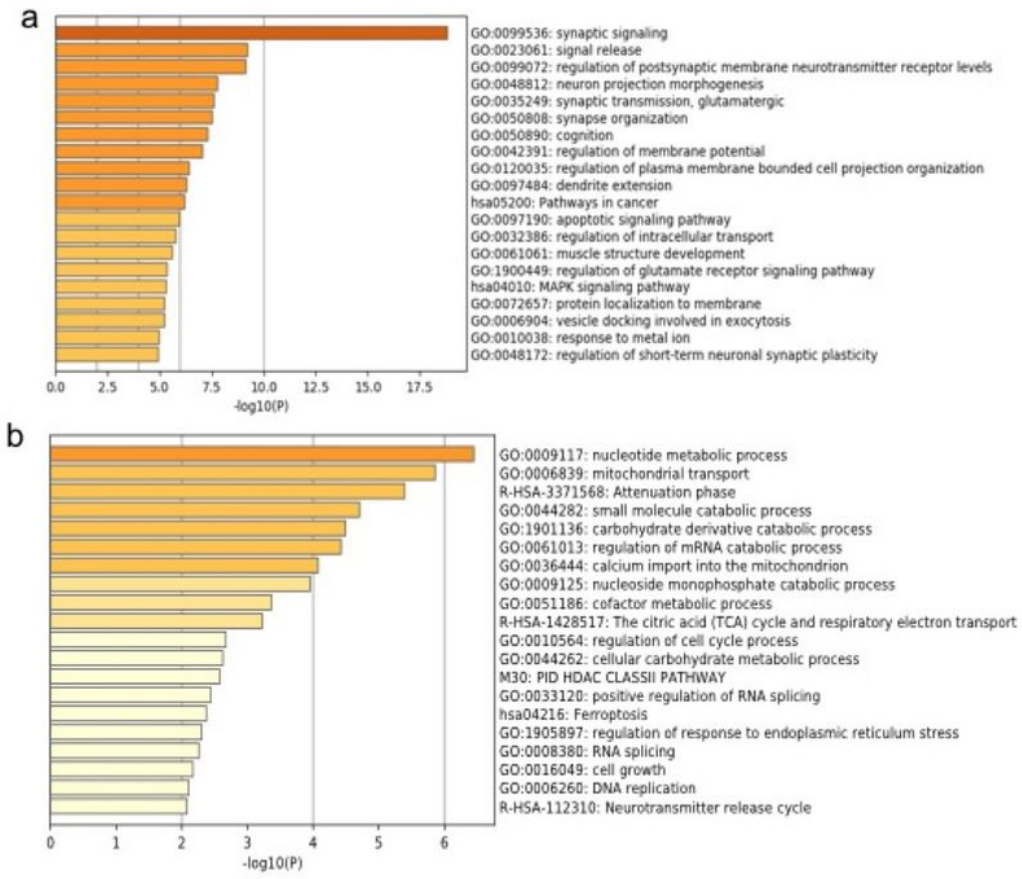

Figure 4

Plot of the enriched GO and KEGG terms in key co-expression modules. a: The plot of enriched biological process and KEEG pathway in the turquoise module. b: The plot of enriched biological process and KEEG pathway in the brown module.