Hub genes of stroke identified by weighted gene co-expression network analysis

Junhong Li (lijh080@163.com)
First Affiliated Hospital of guangxi University of chinese Medicine https://orcid.org/0000-0002-2154-8695

Yang Zhai
Guangxi Inernational Zhuang Medical Hospital

Peng Wu
the First Affiliated Hospital of Guangxi University of Chinese Medicine

Yueqiang Hu
the First Affiliated Hopital of GuangXi University of Chinese Medicine

Wei Chen
the First Affiliated Hospital of GuangXi University of Chinese Medicine

Jinghui Zheng
Ruikang Hospital Affiliated to Guangxi University of Chinese Medicine

Nong Tang
GuangXi University of Chinese Medicine

Research article

Keywords: Stroke, Co-expression, Network analysis, WGCNA, Hub gene

Posted Date: September 20th, 2019

DOI: https://doi.org/10.21203/rs.2.14681/v1

Abstract

Background Microarray-based gene expression profiling has been widely used in biomedical research. Weighted gene co-expression network analysis (WGCNA) can link microarray data directly to clinical traits and to identify rules for predicting pathological stage and prognosis of disease, it has been found useful in many biological processes. Stroke is one of the most common diseases worldwide, yet molecular mechanisms of its pathogenesis are largely unknown. We aimed to construct gene co-expression networks to identify key modules and hub genes associated with the pathogenesis of stroke.Results In this study, we screened out the differentially expressed genes from gene microarray expression profiles, then constructed the free-scale gene co-expression networks to explore the associations between gene sets and clinical features, and to identify key modules and hub genes. Subsequently, functional enrichment and the receiver operating characteristic (ROC) curve analysis were performed. And the results show that a total of 11,747 most variant genes were used for co-expression network construction. Pink and yellow modules were found to be the most significantly related to stroke. Functional enrichment analysis showed that the pink module was mainly involved in regulation of neuron regeneration, and the repair of DNA damage, while the yellow module was mainly enriched in ion transport system dysfunction which were correlated with neuron death. A total of 8 hub genes (PRR11, NEDD9, Notch2, RUNX1-IT1, ANP32A-IT1, ASTN2, SAMHD1 and STIM1) were identified and validated at transcriptional levels (other datasets) and by existing literatures.Conclusions Eight hub genes (PRR11, NEDD9, Notch2, RUNX1-IT1, ANP32A-IT1, ASTN2, SAMHD1 and STIM1) may serve as biomarkers and therapeutic targets for precise diagnosis and treatment of stroke in the future.

Background

Stroke is the second cause of death across the globe, ranking only after ischaemic heart disease and remains one of the leading causes of death globally in the next 15 years [1, 2]. It accounts for 11.3% of all deaths each year of which more than 85% of deaths occur in low- and middle-income countries [3, 4]. Despite declining mortality rates over the past two decades, the global burden of stroke is increasing [5]. According to the current epidemiological data, approximately 16.9 million people were affected by stroke each year[6], and it is estimated that, by 2030, the number of stroke survivors will rise to 77 million[7]. As an archetypical common complex disease mediated by the interplay between genetic and environmental determinants, the proportion of subtypes of stroke vary among populations of different age, race and ethnic origin [4, 5]. This may be one of the complex therapeutic challenge in stroke.

Microarray-based gene expression profiling has been widely used in biomedical research especially in chronic non-communicable diseases areas including neurological diseases[8], cardiovascular disease[9], diabetes[10], cancer[11].In recent literature most of the microarray analysis methods centers on the comparison between the normal and diseased conditions[12]. The identification of differential gene expression is a widely used analytical strategy when screening for potential biomarkers in the diseased state. Many efforts have been focused in the field, but single microarray dataset tends to have high false positive rates and the integration of multiple datasets would significantly increase the reliability and

sensitivity of the findings and reduce those false results. Gene co-expression network analysis, the recent systems biology approach has been increasingly used for describing the correlation patterns among microarray analysis [13-15].

Weighted gene co-expression network analysis (WGCNA) [16, 17], a tool for exploring the system-level functionality of genes, is increasingly used in the gene expression data analysis. It defined co-expression by measuring similarity in gene expression between expression profiles and identified groups of genes that are highly correlated across the samples [17]. WGCNA has been found useful in many biological processes. It can help unravelling the interactions between genes in different modules and hence can be useful to identify candidate biomarkers or therapeutic targets [18, 19]. Besides, WGCNA can link microarray data directly to clinical traits to clarify mechanisms of drug resistance [20], and to identify rules for predicting pathological stage and prognosis of disease [21].

In this study, we aimed to construct co-expression modules using blood samples from stroke. The differentially expressed genes (DEGs) between expression profiles were identified, followed by the biological function, Gene ontology (GO) enrichment analysis. Then, hub genes were identified and the effects of biomarkers for stroke were confirmed by data validation and literature validation. We hoped that these key modules and hub genes might contribute to our understanding of the molecular mechanisms of stroke and might provide the candidate biomarkers for stroke gene therapy.

Results

3.1 Co-expression network construction

The samples of GSE22255 were clustered to detect outliers using hierarchical cluster analysis. One sample was removed and 39 samples were kept (Additional file 1: Figure S1). The left samples were clustered again and clinical traits were converted to color representation and visualized by using heatmap tool package (Fig. 1). In current study, the scale free topology for multiple soft thresholding powers were calculated, and the power of $\beta = 12$ (scale free $R^2 = 0.905$) was selected as an appropriate soft thresholding parameter for weighted co-expression network construction (Fig. 2). A total of 11,747 most variant genes were used for co-expression network construction. Cluster dendrogram and network heatmap of the most variant genes were carried out based on a dissimilarity measure (Fig. 3).

3.2 Key modules identification

After k-means clustering, detecting, calculating and checking module eigengenes, a total of 24 modules were identified (Additional file 1: Table S1). Grey module was not included in any module, so the subsequent analysis was no long performed on grey [22]. Lightgreen (73 genes), pink (317 genes), midnightblue (109 genes) and yellow (870 genes) modules were found to have highly association with status (Fig. 4) (Additional file 1: Figure S2), and this modules were selected as the clinical significant module for further analysis. Since ME was in accordance with the expression level for each module, the lightgreen, pink and midnightblue modules were downregulated, while the yellow module was

upregulated. Moreover, the lightgreen (correlation coefficient r=-0.41, P=0.01), pink(r=-0.33, P=0.04) and midnightblue(r = -0.4, P=0.01) modules were negatively correlated with the disease status. The yellow(r = 0.35, P=0.03) module were positively correlated with the disease status (Fig. 4). In addition, the correlations between gene significance (GS) and the two modules (pink and yellow) were 0.21(P = 0.00) and 0.16(P = 0.00), respectively, which indicated a close relationship with stroke (Fig. 5). So the pink and yellow modules were performed for the subsequent analysis.

3.3 GO enrichment KEGG pathway analysis in two modules

To explore the biological processes and pathways in the pink and yellow module, enrichment analysis was performed. GO enrichment Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were conducted by ClueGO plugin in Cytoscape, and the detailed information was shown in Table1. The result showed that biological processes of the yellow module was mainly enriched in ion transport system dysfunction including potassium ion transport, cellular potassium ion transport and potassium ion transmembrane transport, which were correlated with neuron death. Biological processes of the pink module was mainly involved in regulation of neuron projection regeneration, and the repair of DNA damage such as nucleotide-excision repair, transcription-coupled nucleotide-excision repair which played an important role in prevention of cell death after stroke. Meanwhile, pathways of the yellow module were enriched in neuroactive ligand-receptor interaction, arginine biosynthesis, and alanine, aspartate and glutamate metabolism. In the pink module, pathways were mainly enriched in nucleotide excision repair, retinol metabolism and B cell receptor signaling pathway.

3.4 Hub genes identification and validation

Based the cut-off criteria (|module membership (MM)| > 0.9 and | trait significance (TS)| > 0.2), 11 genes in the yellow module and 13 genes in the pink module were identified as hub genes, respectively. The detailed information was indicated in Table2. Among them, the expression levels of SAM and HD domain containing deoxynucleoside triphosphate triphosphohydrolase 1(SAMHD1) and runt related transcription factor 1- intronic transcript 1(RUNX1-IT1) were significantly lower in stroke samples compared with control samples of GSE22255 dataset (Fig. 6). Moreover, ROC curve indicated that 19 hub genes (all area under curve (AUC)>0.6) exhibited good diagnostic efficiency for control and stroke tissues (Table 3). In order to ensure the robustness and reliability of the results, the ROC analysis of the diagnostic efficiency genes were validated in another two Gene Expression Omnibus (GEO) datasets, GSE16561 and GSE58294. The results showed that 8 key genes with five protein coding genes and six noncoding genes can well distinguish stroke from the control samples (Table 3). The 8 key genes include proline-rich protein 11(PRR11), neural precursor cell expressed, developmentally down-regulated 9(NEDD9), Notch2, RUNX1-IT1, acidic nuclear phosphoprotein 32 family member A- intronic transcript 1(ANP32A-IT1), astrotactin-2(ASTN2), SAMHD1, stromal interaction molecule 1(STIM1).To assess the biological significance of the identified hub genes, we searched the literature and constructed key genes associated network by using Agilent Literature Search plugin in the Cytoscape software (Fig. 7). It contained a total of 50 nodes in the six networks, including 17 nodes with 26 interactions in the PRR11 network, 5 nodes

with 6 interactions in the NEDD9 network, 13 nodes with 28 interactions in the Notch2 network, 4 nodes with 3 interactions in the STIM1 network, 9 nodes with 6 interactions in the SAMHD1 network, and 2 nodes with 1 interactions in the ASTN2 network.

Discussion

Stroke seriously endangers human health, and understanding the underlying molecular mechanisms is critical for precision medicine or personalized medicine. Although the pathogenesis of stroke is extremely complicated, important progress has been made in some areas during the last decades, such as energy metabolism disorders, excitatory amino acid toxicity, penumbra depolarization and apoptosis, which are involved in its pathophysiological processes. However, there is still a way to go to fully elucidate the molecular mechanisms of stroke. Therefore, it is necessary to explore the molecule mechanisms involved in the development of stroke. In this study, we used gene expression datasets from GEO database to construct a weighted gene co-expression network and identified pathways and key genes that are potential biomarkers or therapeutic targets for stroke. We obtained whole genome data of stroke from GEO database and literature for validation as well.

WGCNA was performed to construct free-scale gene co-expression networks, and to explore gene co-expression modules associated with stroke. A total of 11,747 most variant genes were used to construct co-expression network and 24 modules were identified. Pink module with 317 genes and yellow module with 870 genes were found to have significant association with stroke and 26 genes with high connectivity were screened out from the two modules. Among them, 8 hub genes were highly associated with stroke. The hub genes of pink module are PRR11, NEDD9, Notch2, RUNX1-IT1 and ANP32A-IT1, and the hub genes of yellow module are ASTN2, SAMHD1 and STIM1. Co-expression analysis showed that different modules was largely associated with their different functions. As a result, genes in model pink was found to be mainly related to cell apoptosis and neuronal differentiation and genes in model yellow was mainly involved in synaptic form and stroke recovery.

PRR11, which is involved in cell cycle regulation, is related to the regulation of protein-protein interaction and cell signal transduction via Wnt/β-catenin signaling pathway [23]. PRR11 consists of a binary nuclear localization signal, two proline enrichment regions and a zinc finger domain, which regulates gene transcription by binding to duplex DNA [24]. In previous studies, PRR11 was considered to be a disadvantageous factor affecting the regulation of tumor cells, but its role and clinical application value in stroke are largely unknown[24, 25]. Increasing evidence suggests that Notch signaling in cerebral ischemia is involved in inflammation, oxidative stress, apoptosis, angiogenesis, synaptic plasticity and the function of blood-brain barrier [26-28]. Not surprisingly, Notch2 is upregulated with increased cell death shortly after cerebral ischemia injury in hippocampal areas [27]. Meanwhile, a similar increase in Notch2 in apoptotic cells was found after oxygen glucose deprivation treatment [29].In addition, Notch2 signaling is associated with the progression of atherosclerosis [30]. It has been shown that Notch2 mediates quiescence in endothelial cells, while inflammatory cytokines trigger increased Notch2 activity promoting a significant augmentation of apoptosis. It is crucial to understand the Notch2 signaling

dysfunction in stroke for development of new therapeutics that are centred around it [31]. The brain can self-repair by producing new neurons and has the ability to compensate for the loss of function after stroke [32]. NEDD9 (Neuronal precursor cell-expressed, developmentally down-regulated gene) was initially identified in the mouse central nervous system [33, 34]. NEDD9, which is a splicing variant of Cas-L, promotes neurite outgrowth by tyrosine phosphorylated and disappears in the adult brain, is upregulated again for neuronal differentiation after cerebral ischemia [34]. It indicates that NEDD9 has potency for recovery of neurologic function after stroke, and upregulation of NEDD9 may widen the therapeutic time window for cerebral ischemia [34, 35].

SAMHD1, originally identified from a dendritic cell, is a deoxyribonucleoside triphosphate triphosphohydrolase [36-38] and its mutations was recently linked to susceptibility to stroke. Recent studies have revealed that SAMHD1 gene mutations might cause genetic predispositions that interact with other risk factors, resulting in increased vulnerability to stroke [39]. Furthermore, a stroke cohort study reported that SAMHD1 gene mutations may be associated with stroke in the general population [40]. In addition, a functional loss of the SAMHD1 protein resulting from the missense mutations c.64C>T and c.841G>A [41]. In patients with cerebrovascular diseases, lack of SAMHD1 protein expression was associated with decreased expression of IFNB1 and increased expression of IL8 [42].

ASTN2 is a conserved perforin-like membrane protein expressed in the developing and adult brain primarily involved in neuronal development [43, 44]. ASTN2 may modulate a number of protein complexes in neurons that impact synaptic form and regulate synaptic adhesion activity [45]. ASTN2 binds to a number of the interacting proteins, like ROCK2 and SLC12a5, which are implicated in vesicle trafficking and synaptic function. The Ca^{2+} entry is critical for platelet activation and stroke. Glutamate-induced dysregulation of intracellular Ca^{2+} homeostasis is a key mechanism of stroke [46]. STIM1, a Ca^{2+} sensor localized to the endoplasmic reticulum, is involved in regulation of store-operated calcium entry [47]. A defective Ca^{2+} entry mechanism in the mutant platelets activate STIM1, resulting in an impaired platelet aggregate formation [48]. This process could protect the mice from ischemic stroke [49]. Dong M et al reported that STIM1/Orai1 expression was associated with mortality and recurrence in ischaemic stroke patients and was an independent predictor of the 3-month stroke recovery [50].

Intriguingly, there are three hub genes in the present study. Recent studies have shown that long non-coding RNA is considered to be a key regulator of the pathogenesis of stroke [51]. Two key noncoding RNAs (RUNX1-IT1 and ANP32A-IT1) were identified as diagnostic efficiency genes in present study. LncRNA RUNX1-IT1, and ANP32A-IT1 are the intronic transcript 1 from their respective genes [52, 53]. RUNX1 IT1 plays a tumour suppressive role and is downregulated in colorectal cancer by regulating cell proliferation, migration, and apoptosis [53]. Although the functions of the RUNX1-IT1 have been identified in different diseases [54], its biological roles in stroke still remains unknown. ANP32A has multiple functions involved in neurons differentiation, brain development, and neuritogenesis. Modulating ANP32A signaling could help manage oxidative stress in brain [55] and restore cognitive function [56] with therapeutic implications for neurological disease. To our best knowledge, no study focusing on the function of RUNX1-IT1 and ANP32A-IT1 has been published in stroke until now.

Since the neuronal death in cerebral ischemic area of stroke was accompanied by the abnormal genes expression, the hub genes of pink module enriched the biological process could possibly regulate the cellular signals transduction involved in inflammation, oxidative stress, apoptosis, angiogenesis, and synaptic plasticity. While the hub genes of yellow module might took part in synaptic form, antiplatelet aggregate formation and short-term prognosis after stroke. Therefore, we speculated that pink module and yellow module were regarded as critical modules in neuronal death, new synapse formation, and functional recovery of stroke. What's more, non-coding genes subset in the hub genes was identified in the present study. Considering that non-coding RNAs play a key regulator in the pathogenesis of stroke, further studies involving non-coding RNAs populations and exploring the mechanism underlying the effect of them in stroke will be valuable.

Co-expression analyses revealed the mRNA regulation expression network in stroke. In the present study, we used WGCNA to construct gene co-expression networks and to explore the relationships between modules and clinical traits. Due to the high degree of consistency in the expression relationships of the module, the genes of the same module might share a common biological role. Our results provided valuable tip for the basic and clinical research of stroke. Although there are similar limitations with most other data mining approaches [57], we adopt corresponding methods to reduce possible bias. In order to increase the credibility of WGCNA results, we used matched stroke and normal samples for analysis, and used external data from GEO database to validate the results.

Conclusions

The hub genes PRR11, NEDD9, Notch2, RUNX1-IT1, ANP32A-IT1, ASTN2, SAMHD1 and STIM1 were found to be significantly correlated with stroke, and were verified in another two datasets. These hub genes may play as the potential diagnostic and prognostic biomarkers of stroke, which also needed much further research. Moreover, pink and yellow module were considered as critical modules in neuronal death, new synapse formation, and functional recovery of stroke.

Methods

Microarray data

The gene expression profile of GSE22255 [58], GSE16561 [59] and GSE58294 [60] were obtained from GEO database (https://www.ncbi.nlm.nih.gov/geo/). GSE22255 contains 20 stroke samples and 20 nonstroke control samples. As for the GSE58294, it contains a total of 92 blood samples including 69 stroke samples and 23 control samples. GSE16561 included 39 ischemic stroke samples and 24 healthy control samples. Raw data in this study were based on the Affymetrix platform GPL570 and Illumina platform GPL6883.

Data preprocessing and differentially expressed genes analysis

A data analysis workflow of this study was shown in Fig. 8. Analysis was performed using the R software (version 3.5.1). Raw data were processed using a Robust Multi-array Average (RMA) algorithm based on a precompiled C language function in limma package which including background correcting, normalizing and calculating expression[61]. Missing values in the raw data were imputed using the knn function in the impute package, and any probe absent from all CEL files was eliminated. Then, the probesets were annotated by using annotate package. Next, we used the built-in match function in R to match probesets to their gene symbol. If multiple probes match a single gene, the probes with the highest interquartile range (IQR) were selected as described in previous studies [62]. Following the expression matrix with 23,495 genes was generated for the subsequent analysis. The top 50% most variant genes (11,747 genes) [63] were considered to be DEGs and selected for WGCNA analysis.

WGCNA analysis

The co-expression network analysis of expression data from GSE22255 was conducted using a convenient one-step network construction method in the WGCNA package [16] to find modules related to stroke. First, in order to ensure the reliability of expression data, samples were clustered to detect outliers. Second, the thresholding power β based on the criterion of approximate scale-free topology was selected for constructing a weighted gene network. The soft threshold calculates adjacency which ranges from 0 to 1, so that the constructed network conforms to the power-law distribution and is closer to the real biological network state [22]. Third, the scale-free gene network was constructed and genes with similar patterns of expression (modules) was identified using blockwiseModules function in the WGCNA package. This function uses a dynamic tree-cutting algorithm to cut the hierarchical clustering tree into branches and [64]. WGCNA approach not only takes the association between the two connected genes into account, but also considers the topological overlap measure which representing the overlap in shared adjacent genes. Based on the single block and block-wise module colors, we then calculated the module eigengene (ME) which represented the expression level for each module. Next, to identify the clinical significant module, we calculated the strength of interaction between clinical trait and ME in each module. Finally, the GS in the module was defined as average p-value of each gene, and the module significance represented the average GS of all the genes in a given module [63].

Enrichment and modules network analysis

The ClueGO plugin in Cytoscape (version 3.7.0) (https://cytoscape.org/) [65] was used to perform GO and KEGG pathway enrichment analysis of clinical significant modules [66]. The identification of significant GO terms and KEGG pathways with $P<0.05$ were selected. After clinical significant modules was identified, the MCODE plugin in Cytoscape was used to construct modules network and sub-networks were extracted for further analysis.

Hub gene identification and validation

Hub gene has been considered as functionally significant. The hub gene in a module was identified by module connectivity when the absolute value of the Pearson's correlation of MM is larger than 0.9[63, 67].

If the absolute value of the Pearson's correlation of TS is more than 0.2[63], a gene was considered to be highly correlated with a certain clinical trait. Furthermore, the genes with both |MM|≥0.9 and |TS|≥0.2 were regarded as “real” hub genes. In the verification set of GSE16561 and GSE58294, the ROC curve analysis were performed to validate the expression of hub genes in stroke and normal samples. Moreover, the AUC was calculated by using pROCR package [68]. Larger AUC value of a gene indicated that it can well distinguish stroke from the control samples, and the hub gene of AUC > 0.6 in the three datasets was defined as a diagnostic efficiency gene [22, 69].

Abbreviations

WGCNA: Weighted gene co-expression network analysis; ROC: receiver operating characteristic; DEGs: differentially expressed genes; GO: Gene ontology; GS: gene significance; KEGG: Kyoto Encyclopedia of Genes and Genomes; MM: module membership; TS: trait significance; SAMHD1: SAM and HD domain containing deoxynucleoside triphosphate triphosphohydrolase 1; RUNX1-IT1: runt related transcription factor 1- intronic transcript 1; AUC: area under curve; GEO: Gene Expression Omnibus; PRR11: proline-rich protein 11; NEDD9: neural precursor cell expressed, developmentally down-regulated 9; ANP32A-IT1: acidic nuclear phosphoprotein 32 family member A- intronic transcript 1; ASTN2: astrotactin-2; STIM1: stromal interaction molecule 1;RMA: Robust Multi-array Average; IQR: interquartile range; ME: module eigengene.

Declarations

Acknowledgements

Not applicable.

Authors’ contributions

JL and NT conceived and designed the study; YZ, PW and YH carried out experiments and conducted data analysis; WC and JZ wrote the manuscript. All authors have read and approved the manuscript.

Funding

The work was supported by the Basic Ability Improvement Project for Middle-aged and Young Teachers in Colleges and Universities in Guangxi (2018KY0284 and 2018KY0294), Guangxi University of Chinese Medicine Scientific Research Project (2017QN021), Guangxi University of Chinese Medicine Qihuang Engineering High-level Talent Team Cultivation Project (No.2018003), Guangxi Science and Technology Project(guike AB16380324).

Availability of data and materials

All data generated or analysed during this study are included in this published article.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Department of Encephalopathy, The First Affiliated Hospital of Guangxi University of Chinese Medicine, Guangxi, 530023, China. [2] Department of Neurology, Guangxi International Zhuang Medical Hospital, Guangxi, 530023, China. [3]Department of Cardiovascular Diseases, Ruikang Hospital Affiliated to Guangxi University of Chinese Medicine, Guangxi, 530011, China. [4]Guangxi Key Laboratory of Chinese Medicine Foundation Research, Guangxi University of Chinese Medicine, Guangxi, 530000, China. [5]Principal's Office, Guangxi University of Chinese Medicine, Nanning, Guangxi, 530000, China.

References

1. Organization WH. The Top 10 Causes of Death [http://www.who.int/en/news-room/fact-sheets/detail/the-top-10-causes-of-death]. Accessed May 24 2018.
2. Stinear CM. Prediction of motor recovery after stroke: advances in biomarkers. Lancet Neurol. 2017; 16(10):826-36.
3. Limbole EB, Magne J, Lacroix P. Stroke characterization in Sun Saharan Africa: Congolese population. Int J Cardiol. 2017; 240:392-7.

4. Writing Group M, Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, Das SR, de Ferranti S, Despres JP et al. Heart Disease and Stroke Statistics-2016 Update: A Report From the American Heart Association. Circulation. 2016; 133(4):e38-360.
5. Hankey GJ. Stroke. Lancet. 2017; 389(10069):641-54.
6. Feigin VL, Forouzanfar MH, Krishnamurthi R, Mensah GA, Connor M, Bennett DA, Moran AE, Sacco RL, Anderson L, Truelsen T, et al. Global and regional burden of stroke during 1990-2010: findings from the Global Burden of Disease Study 2010. Lancet. 2014;383(9913):245-54.
7. Bejot Y, Daubail B, Giroud M. Epidemiology of stroke and transient ischemic attacks: Current knowledge and perspectives. Rev Neurol (Paris). 2016; 172(1):59-68.
8. Nakaoka H, Tajima A, Yoneyama T, Hosomichi K, Kasuya H, Mizutani T, Inoue I. Gene expression profiling reveals distinct molecular signatures associated with the rupture of intracranial aneurysm. Stroke. 2014; 45(8):2239-45.
9. Steenman M, Lamirault G, Le Meur N, Leger JJ. Gene expression profiling in human cardiovascular disease. Clin Chem Lab Med. 2005; 43(7):696-701.
10. Chen J, Meng Y, Zhou J, Zhuo M, Ling F, Zhang Y, Du H, Wang X. Identifying candidate genes for Type 2 Diabetes Mellitus and obesity through gene expression profiling in multiple tissues or cells. J Diabetes Res. 2013; 2013:970435.
11. Rojo F, Domingo L, Sala M, Zazo S, Chamizo C, Menendez S, Arpi O, Corominas JM, Bragado R, Servitja S et al. Gene expression profiling in true interval breast cancer reveals overactivation of the mTOR signaling pathway. Cancer Epidemiol Biomarkers Prev. 2014; 23(2):288-99.
12. Yip L, Fuhlbrigge R, Atkinson MA, Fathman CG. Impact of blood collection and processing on peripheral blood gene expression profiling in type 1 diabetes. BMC Genomics. 2017; 18(1):636.
13. Liang JW, Fang ZY, Huang Y, Liuyang ZY, Zhang XL, Wang JL, Wei H, Wang JZ, Wang XC, Zeng J, et al. Application of Weighted Gene Co-Expression Network Analysis to Explore the Key Genes in Alzheimer's Disease. J Alzheimers Dis. 2018; 65(4):1353-64.
14. Shi H, Zhang L, Qu Y, Hou L, Wang L, Zheng M. Prognostic genes of breast cancer revealed by gene co-expression network analysis. Oncol Lett. 2017; 14(4):4535-42.
15. Chen JA. Gene Co-Expression Network Analysis Implicates microRNA Processing in Parkinson's Disease Pathogenesis. Neurodegener Dis. 2018; 18(4):191-9.
16. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008; 9:559.
17. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol. 2005; 4:Article17.
18. Giulietti M, Occhipinti G, Righetti A, Bracci M, Conti A, Ruzzo A, Cerigioni E, Cacciamani T, Principato G, Piva F. Emerging Biomarkers in Bladder Cancer Identified by Network Analysis of Transcriptomic Data. Front Oncol. 2018; 8:450.

19. Costa RL, Boroni M, Soares MA. Distinct co-expression networks using multi-omic data reveal novel interventional targets in HPV-positive and negative head-and-neck squamous cell cancer. Sci Rep. 2018; 8(1):15254.
20. Chen Y, Bi F, An Y, Yang Q. Coexpression network analysis identified Kruppel-like factor 6 (KLF6) association with chemosensitivity in ovarian cancer. J Cell Biochem. 2018.
21. Chen L, Yuan L, Qian K, Qian G, Zhu Y, Wu CL, Dan HC, Xiao Y, Wang X. Identification of Biomarkers Associated With Pathological Stage and Prognosis of Clear Cell Renal Cell Carcinoma by Co-expression Network Analysis. Front Physiol. 2018; 9:399.
22. Yang Q, Wang R, Wei B, Peng C, Wang L, Hu G, Kong D, Du C. Candidate Biomarkers and Molecular Mechanism Investigation for Glioblastoma Multiforme Utilizing WGCNA. Biomed Res Int. 2018; 2018:4246703.
23. Yang P, Qiu Z, Jiang Y, Dong L, Yang W, Gu C, Li G, Zhu Y. Silencing of cZNF292 circular RNA suppresses human glioma tube formation via the Wnt/beta-catenin signaling pathway. Oncotarget. 2016;**7**(39):63449-55.
24. Qiao W, Wang H, Zhang X, Luo K. Proline-rich protein 11 silencing inhibits hepatocellular carcinoma growth and epithelial-mesenchymal transition through beta-catenin signaling. Gene. 2019; 681:7-14.
25. Zhu J, Hu H, Wang J, Yang Y, Yi P. PRR11 Overexpression Facilitates Ovarian Carcinoma Cell Proliferation, Migration, and Invasion Through Activation of the PI3K/AKT/beta-Catenin Pathway. Cell Physiol Biochem. 2018; 49(2):696-705.
26. Cai Z, Zhao B, Deng Y, Shangguan S, Zhou F, Zhou W, Li X, Li Y, Chen G. Notch signaling in cerebrovascular diseases (Review). Mol Med Rep. 2016; 14(4):2883-98.
27. Arumugam TV, Baik SH, Balaganapathy P, Sobey CG, Mattson MP, Jo DG. Notch signaling and neuronal death in stroke. Prog Neurobiol. 2018; 165-167:103-16.
28. Peterson SM, Turner JE, Harrington A, Davis-Knowlton J, Lindner V, Gridley T, Vary CPH, Liaw L. Notch2 and Proteomic Signatures in Mouse Neointimal Lesion Formation. Arterioscler Thromb Vasc Biol. 2018; 38(7):1576-93.
29. Alberi L, Chi Z, Kadam SD, Mulholland JD, Dawson VL, Gaiano N, Comi AM. Neonatal stroke in mice causes long-term changes in neuronal Notch-2 expression that may contribute to prolonged injury. Stroke. 2010; 41(10 Suppl):S64-71.
30. Davis-Knowlton J, Turner JE, Turner A, Damian-Loring S, Hagler N, Henderson T, Emery IF, Bond K, Duarte CW, Vary CPH, et al. Characterization of smooth muscle cells from human atherosclerotic lesions and their responses to Notch signaling. Lab Invest. 2019; 99(3):290-304.
31. Ables JL, Breunig JJ, Eisch AJ, Rakic P. Not(ch) just development: Notch signalling in the adult brain. Nat Rev Neurosci. 2011; 12(5):269-83.
32. Sun Y, Sun X, Qu H, Zhao S, Xiao T, Zhao C. Neuroplasticity and behavioral effects of fluoxetine after experimental stroke. Restor Neurol Neurosci. 2017; 35(5):457-68.
33. Lindvall O, Kokaia Z. Neurogenesis following Stroke Affecting the Adult Brain. Cold Spring Harb Perspect Biol. 2015; **7**(11).

34. Sasaki T, Iwata S, Okano HJ, Urasaki Y, Hamada J, Tanaka H, Dang NH, Okano H, Morimoto C. Nedd9 protein, a Cas-L homologue, is upregulated after transient global ischemia in rats: possible involvement of Nedd9 in the differentiation of neurons after ischemia. Stroke. 2005; 36(11):2457-62.
35. Shagisultanova E, Gaponova AV, Gabbasov R, Nicolas E, Golemis EA. Preclinical and clinical studies of the NEDD9 scaffold protein in cancer and other diseases. Gene . 2015; 567(1):1-11.
36. Goldstone DC, Ennis-Adeniran V, Hedden JJ, Groom HC, Rice GI, Christodoulou E, Walker PA, Kelly G, Haire LF, Yap MW, et al. HIV-1 restriction factor SAMHD1 is a deoxynucleoside triphosphate triphosphohydrolase. Nature. 2011; 480(7377):379-82.
37. Hrecka K, Hao C, Gierszewska M, Swanson SK, Kesik-Brodacka M, Srivastava S, Florens L, Washburn MP, Skowronski J. Vpx relieves inhibition of HIV-1 infection of macrophages mediated by the SAMHD1 protein. Nature. 2011; 474(7353):658-61.
38. Li N, Zhang W, Cao X. Identification of human homologue of mouse IFN-gamma induced protein from human dendritic cells. Immunol Lett. 2000;74(3):221-4.
39. Ji X, Wu Y, Yan J, Mehrens J, Yang H, DeLucia M, Hao C, Gronenborn AM, Skowronski J, Ahn J, et al. Mechanism of allosteric activation of SAMHD1 by dGTP. Nat Struct Mol Biol. 2013; 20(11):1304-9.
40. Li W, Xin B, Yan J, Wu Y, Hu B, Liu L, Wang Y, Ahn J, Skowronski J, Zhang Z, et al. SAMHD1 Gene Mutations Are Associated with Cerebral Large-Artery Atherosclerosis. Biomed Res Int. 2015;2015:739586.
41. Xin B, Jones S, Puffenberger EG, Hinze C, Bright A, Tan H, Zhou A, Wu G, Vargus-Adams J, Agamanolis D, et al. Homozygous mutation in SAMHD1 gene causes cerebral vasculopathy and early onset stroke. Proc Natl Acad Sci U S A. 2011; 108(13):5372-7.
42. Holger T, Marcel DM, Katarzyna B, Christel G, Wolfram S, Gudrun N, Michael F, Gerhard K, Johannes R, Peter NJHM. Cerebral arterial stenoses and stroke: novel features of Aicardi-Goutières syndrome caused by the Arg164X mutation in SAMHD1 are associated with altered cytokine expression. 2010; 31(11):E1836-50.
43. Wilson PM, Fryer RH, Fang Y, Hatten ME. Astn2, a novel member of the astrotactin gene family, regulates the trafficking of ASTN1 during glial-guided neuronal migration. J Neurosci. 2010; 30(25):8529-40.
44. Ni T, Harlos K, Gilbert RJOB. Structure of astrotactin-2: a conserved vertebrate-specific and perforin-like membrane protein involved in neuronal development. Open Biol.2016; 6(5):160053.
45. Behesti H, Fore TR, Wu P, Horn Z, Leppert M, Hull C, Hatten ME. ASTN2 modulates synaptic strength by trafficking and degradation of surface proteins. Proc Natl Acad Sci U S A. 2018; 115(41):E9717-26.
46. Song Q, Gou WL, Zou YL. FAM3A Protects Against Glutamate-Induced Toxicity by Preserving Calcium Homeostasis in Differentiated PC12 Cells. Cell Physiol Biochem. 2017; 44(5):2029-41.
47. Wolfgang B, Chauhan AK, Wagner DD. Glycoprotein Ibalpha and von Willebrand factor in primary platelet adhesion and thrombus formation: lessons from mutant mice. Thromb Haemost.2008; 100(02):264-70.

48. Ohara H, Nabika TJBBRC. A nonsense mutation of Stim1 identified in stroke-prone spontaneously hypertensive rats decreased the store-operated calcium entry in astrocytes. Biochem Biophys Res Commun. 2016; 476(4):406-11.
49. Gotru SK, Chen W, Kraft P, Becker IC, Wolf K, Stritt S, Zierler S, Hermanns HM, Rao D, Perraud ALJA. TRPM7 Kinase Controls Calcium Responses in Arterial Thrombosis and Stroke in Mice. Arterioscler Thromb Vasc Biol.2018; 38(2):344-52.
50. Dong M, Zheng N, Ren LJ, Zhou H, Liu J. Increased expression of STIM1/Orai1 in platelets of stroke patients predictive of poor outcomes. Eur J Neurol. 2017; 24(7):912-9.
51. Yang L, Wang H, Shen Q, Feng L, Jin H. Long non-coding RNAs involved in autophagy regulation. Cell Death Dis.2017; 8(10):e3073.
52. Ramsey J, Butnor K, Peng Z, Leclair T, van der Velden J, Stein G, Lian J, Kinsey CM. Loss of RUNX1 is associated with aggressive lung adenocarcinomas. J Cell Physiol. 2018; 233(4):3487-97.
53. Shi J, Zhong X, Song Y, Wu Z, Gao P, Zhao J, Sun J, Wang J, Liu J, Wang Z. Long non-coding RNA RUNX1-IT1 plays a tumour-suppressive role in colorectal cancer by inhibiting cell proliferation and migration. Cell Biochem Funct. 2019; 37(1):11-20.
54. Yang K, Hou Y, Li A, Li Z, Wang W, Xie H, Rong Z, Lou G, Li K. Identification of a six-lncRNA signature associated with recurrence of ovarian cancer. Sci Rep.2017; **7**(1):752.
55. Cornelis FMF, Monteagudo S, Guns LKA, den Hollander W, Nelissen R, Storms L, Peeters T, Jonkers I, Meulenbelt I, Lories RJ. ANP32A regulates ATM expression and prevents oxidative stress in cartilage, brain, and bone. Sci Transl Med. 2018; 10(458).
56. Chai GS, Feng Q, Ma RH, Qian XH, Luo DJ, Wang ZH, Hu Y, Sun DS, Zhang JF, Li XJ, et al. Inhibition of Histone Acetylation by ANP32A Induces Memory Deficits. J Alzheimers Dis.2018; 63(4):1537-46.
57. Wan Q, Tang J, Han Y, Wang D. Co-expression modules construction by WGCNA and identify potential prognostic markers of uveal melanoma. Exp Eye Res. 2018; 166:13-20.
58. Krug T, Gabriel JP, Taipa R, Fonseca BV, Domingues-Montanari S, Fernandez-Cadenas I, Manso H, Gouveia LO, Sobral J, Albergaria I et al. TTC7B emerges as a novel risk factor for ischemic stroke through the convergence of several genome-wide approaches. J Cereb Blood Flow Metab. 2012; 32(6):1061-72.
59. Barr TL, Conley Y, Ding J, Dillman A, Warach S, Singleton A, Matarin M. Genomic biomarkers and cellular pathways of ischemic stroke by RNA gene expression profiling. Neurology. 2010; **75**(11):1009-14.
60. Stamova B, Jickling GC, Ander BP, Zhan X, Liu D, Turner R, Ho C, Khoury JC, Bushnell C, Pancioli A, et al. Gene expression in peripheral immune cells following cardioembolic stroke is sexually dimorphic. PLoS One. 2014; 9(7):e102550.
61. Ritchie ME, Belinda P, Di W, Yifang H, Law CW, Wei S, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res. 2015; 43(7):e47.
62. Zhang X, Feng H, Li Z, Li D, Liu S, Huang H, Li M. Application of weighted gene co-expression network analysis to identify key modules and hub genes in oral squamous cell carcinoma tumorigenesis.

Onco Targets Ther. 2018; 11:6001-21.
63. Tang J, Kong D, Cui Q, Wang K, Zhang D, Gong Y, Wu G. Prognostic Genes of Breast Cancer Identified by Gene Co-expression Network Analysis. Front Oncol. 2018; 8:374.
64. Langfelder P, Zhang B, Horvath S. Defining clusters from a hierarchical cluster tree: the Dynamic Tree Cut package for R. Bioinformatics. 2008; 24(5):719-20.
65. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003; 13(11):2498-504.
66. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012; 16(5):284-7.
67. Yuan L, Chen L, Qian K, Qian G, Wu CL, Wang X, Xiao Y. Co-expression network analysis identified six hub genes in association with progression and prognosis in human clear cell renal cell carcinoma (ccRCC). Genom Data. 2017;14:132-40.
68. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011; 12:77.
69. Long F, Su JH, Liang B, Su LL, Jiang SJ. Identification of Gene Biomarkers for Distinguishing Small-Cell Lung Cancer from Non-Small-Cell Lung Cancer Using a Network-Based Approach. Biomed Res Int. 2015; 2015:685303.

Tables

Table1 GO enrichment analysis of yellow and pink module (top 4 in each module were listed)

Module	GO ID	GO Term	P value	Associated Genes
yellow	GO:0030001	metal ion transport	0.0013	ABCC9, AGXT, ANK2, APLNR, AQP1, ASIC1, ATP12A, ATP13A5, ATP7B, CACNA1H, CACNG6, CASQ2, CMPK1, CYP27B1, DMD, DPP6, EPPIN, F2RL3, FGF2, GAL, GRIN2B, GRM6, KCNA5, KCNG4, KCNJ1, KCNJ3, KCNV1, NCOA6, NETO1, NMUR2, NOS3, PANX1, PARP3, PKD2, PPOX, SLC12A5, SLC13A3, SLC17A2, SLC23A1, SLC39A9, SLC41A2, SLC4A8, SLC5A7, SLC9A8, SNAP91, TRPA1, UCN
yellow	GO:0098660	inorganic ion transmembrane transport	0.0169	ABCC9, ANK2, APLNR, AQP1, ASIC1, ATP12A, ATP13A5, ATP7B, CACNA1H, CACNG6, CASQ2, CFTR, CLCC1, CLDN4, CMPK1, DMD, DPP6, F2RL3, FGF2, G6PC, GABRG1, GAL, GRIN2B, KCNA5, KCNG4, KCNJ1, KCNJ3, KCNV1, NCOA6, NETO1, PANX1, PARP3, PKD2, SLC12A5, SLC17A2, SLC23A1, SLC41A2, SLC9A8, SNAP91, TRPA1
yellow	GO:0009617	response to bacterium	0.0179	ABCC11, ADIPOQ, ALAD, ARHGEF28, ASS1, C10orf99, CHST4, CMPK1, CPS1, CRP, CYP1A2, CYP27B1, DCN, DEFB106A, EPPIN, FMO1, GKN2, HERC6, KYAT1, LIAS, LYPD8, MBL2, MRO, NLRP1, NOS3, PLSCR4, PNLIPRP2, PRL, RARA, RNASEH2A, S100A7A, SPAG11A, TINAGL1, TRAPPC2, UGT1A1
yellow	GO:0010035	response to inorganic substance	0.0080	ABCC11, ALAD, AQP1, ASS1, ATP7B, BAD, CACNA1H, CASQ2, CPO, CPS1, CYP1A2, DMD, EGFR, F2RL3, FXN, HBB, KCNA5, NCOA6, NOS3, PKD2, PON1, RASGRP2, RNASEH2A, S100A16, SLC17A2, SNAP91, TAT, TFAP2A, TINAGL1, TNNT2, TRAF2, TRPA1, TSHB
pink	GO:0000186	activation of MAPKK activity	0.0001	JAK2, KIDINS220, MAP4K4, RAF1, TAOK1, ZHX2
pink	GO:0006289	nucleotide-excision repair	0.0021	CUL4A, GTF2H3, POLK, POLR2J, RPA4, TCEA1
pink	GO:0031102	neuron projection regeneration	0.0004	ATF7IP, JAK2, MAP4K4, RTN4, TEP1
pink	GO:0006283	transcription-coupled nucleotide-excision repair	0.0023	CUL4A, GTF2H3, POLK, POLR2J, TCEA1
yellow	KEGG:04080	Neuroactive ligand-receptor interaction	0.0119	APLNR, CGA, F2RL3, GABRG1, GRIN2B, GRM6, HTR1E, LEPR, LHCGR, NMUR2, PRL, TAAR2, TAAR9, TRHR, TSHB
yellow	KEGG:00250	Alanine, aspartate and glutamate metabolism	0.0142	AGXT, AGXT2, ASS1, CPS1
yellow	KEGG:00220	Arginine biosynthesis	0.0181	ASS1, CPS1, NOS3
yellow	KEGG:05410	Hypertrophic cardiomyopathy (HCM)	0.0246	CACNG6, DMD, ITGA10, ITGB5, TNNT2, TPM4
pink	KEGG:03440	Homologous recombination	0.0045	ATM, RPA4, XRCC2
pink	KEGG:03420	Nucleotide	0.0066	CUL4A, GTF2H3, RPA4

pink	KEGG:00830	excision repair Retinol metabolism	0.0174	ADH4, CYP3A4, RDH11
pink	KEGG:04662	B cell receptor signaling pathway	0.0203	DAPP1, INPP5D, RAF1

GO gene-ontology, *BP* biological process, *KEGG* Kyoto Encyclopedia of Genes and Genomes

Table2 Hub genes in the yellow and pink modules

Module Color	Gene Symbol	MM	p-value of MM
yellow	SAMHD1	0.9212	0.0000
yellow	POLK	0.9171	0.0000
yellow	LINC01192	0.9098	0.0000
yellow	NEGR1-IT1	0.9315	0.0000
yellow	LOC101927815	0.9120	0.0000
yellow	INO80C	0.9025	0.0000
yellow	CACNG6	0.9135	0.0000
yellow	ASTN2	0.9332	0.0000
yellow	ANKRD20A1	0.9357	0.0000
yellow	CCDC169	0.9400	0.0000
yellow	ZNF33A	0.9024	0.0000
pink	FAM49B	0.9100	0.0000
pink	N4BP2L2	0.9430	0.0000
pink	SUPT20H	0.9212	0.0000
pink	MACF1	0.9371	0.0000
pink	C4orf29	0.9223	0.0000
pink	NEDD9	0.9256	0.0000
pink	Notch2	0.9102	0.0000
pink	KDM4C	0.9236	0.0000
pink	PRR11	0.9116	0.0000
pink	VTI1A	0.9256	0.0000
pink	RUNX1-IT1	0.9230	0.0000
pink	FOXP1-IT1	0.9100	0.0000
pink	ANP32A-IT1	0.9024	0.0000

MM Pearson's correlation of gene-module membership

Table3 AUC of hub genes in three datasets

Hub genes	Module	Diagnostic efficiency	GSE22255	GSE16561	GSE58294
FAM49B	pink	NO	0.698	0.628	0.473
VTI1A	pink	NO	0.608	-	0.544
RUNX1-IT1	pink	YES	0.743	-	0.969
PRR11	pink	YES	0.645	0.634	0.885
ANP32A-IT1	pink	YES	0.665	-	0.743
KDM4C	pink	YES	0.723	0.553	0.831
N4BP2L2	pink	YES	0.783	0.685	0.576
SUPT20H	pink	YES	0.718	0.597	0.677
MACF1	pink	YES	0.618	0.449	0.883
FOXP1-IT1	pink	YES	0.648	0.534	0.71
NEDD9	pink	YES	0.628	0.612	0.645
NOTCH2	pink	YES	0.651	-	0.601
C4orf29	pink	YES	0.705	-	0.543
CACNG6	yellow	NO	0.652	0.482	0.649
INO80C	yellow	NO	0.665	0.535	0.507
LOC101927815	yellow	NO	0.568	-	0.535
ASTN2	yellow	YES	0.635	0.618	0.955
SAMHD1	yellow	YES	0.627	-	0.842
POLK	yellow	YES	0.652	-	0.805
STIM1	yellow	YES	0.655	-	0.675
CCDC169	yellow	YES	0.583	-	0.7
NEGR1-IT1	yellow	YES	0.6	-	0.678
ANKRD20A1	yellow	YES	0.52	0.592	0.791
ZNF33A	yellow	YES	0.547	0.656	0.698

Figures

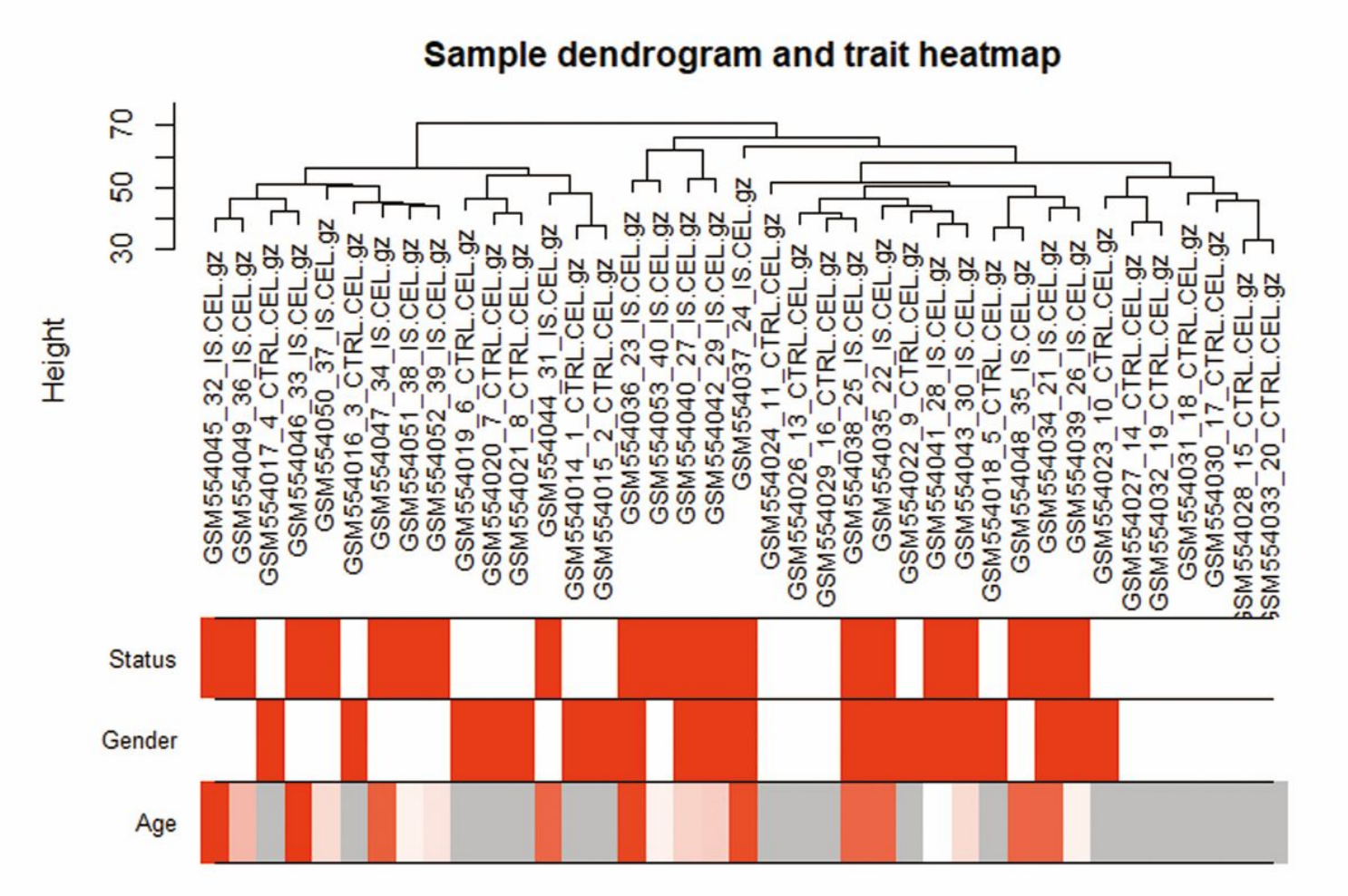

Figure 1

Sample dendrogram and trait heatmap. Re-cluster 39 samples and white means low, red means high and grey means missing entry in the trait heatmap

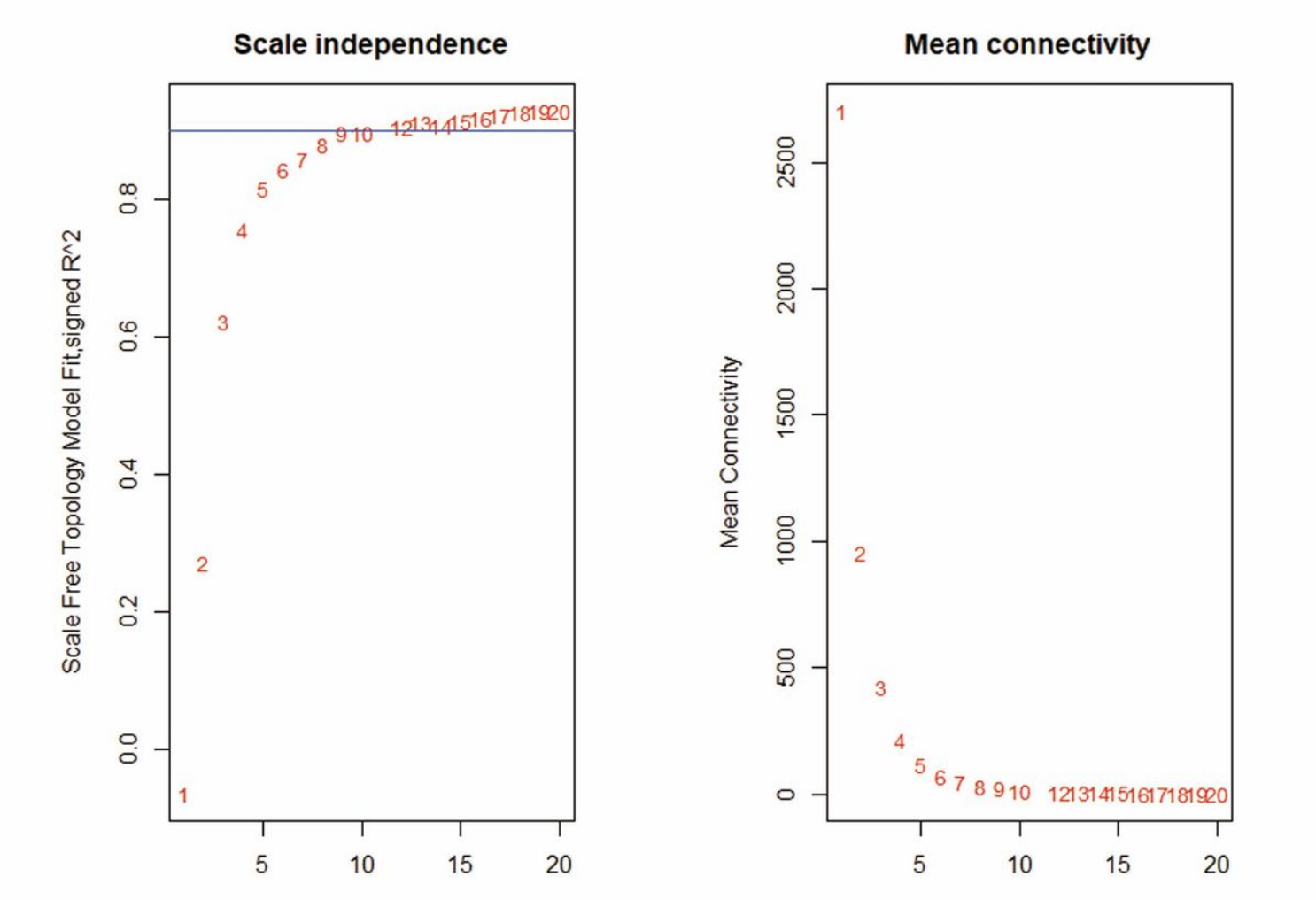

Figure 2

Various soft thresholding powers. The left panel shows the scale-free topology model fit index as a function of the soft-thresholding power. The right panel displays the mean connectivity (degree) as a function of the soft-thresholding power

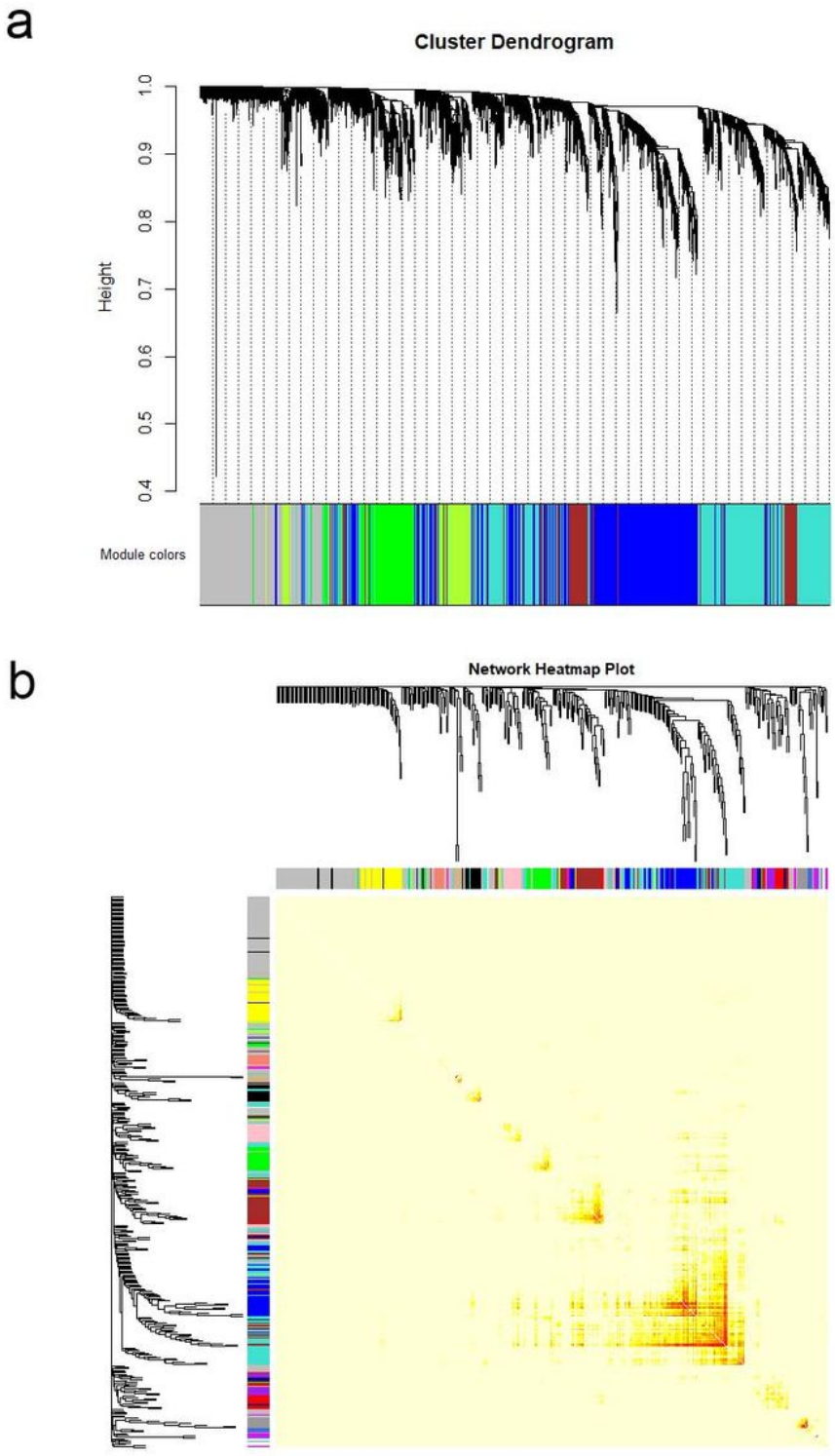

Figure 3

Cluster dendrogram and network heatmap plot. a Dendograms produced by dissimilarity based on topological overlaps with assigned module colors. The degree of gene conservation in the datasets are represented by the same module colours. b Visualizing the gene network heatmap plot. The network heatmap depicts the topological overlap matrix between among all genes. Yellow color represents low

overlap and darker red color represents higher overlap. Genotype maps and module assignments are also shown on the left and top

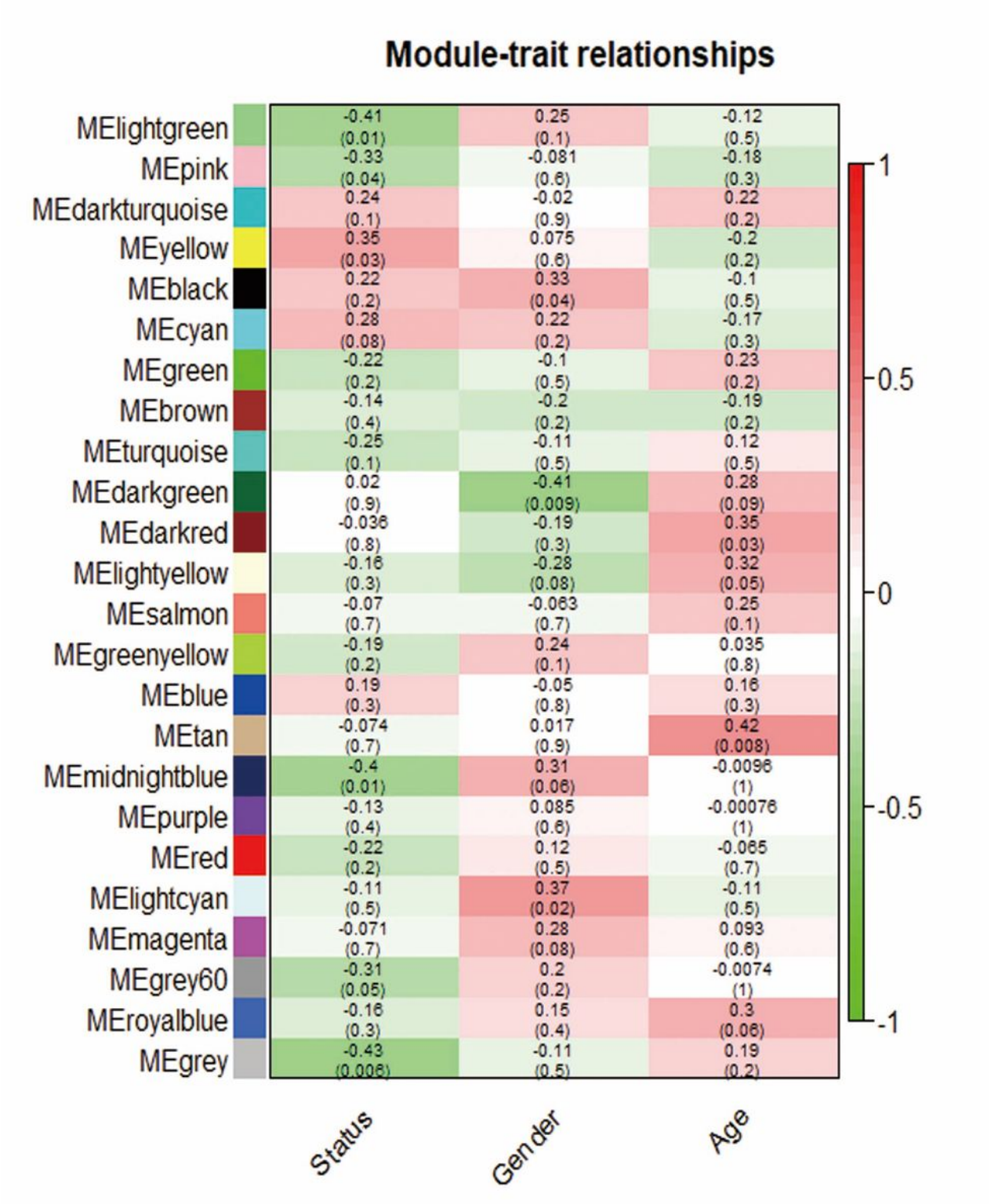

Figure 4

Module-trait relationships. Each row corresponds to a module eigengene, column to a trait. Each cell contains the corresponding correlation and p-value. The table is color-coded by correlation according to the color legend

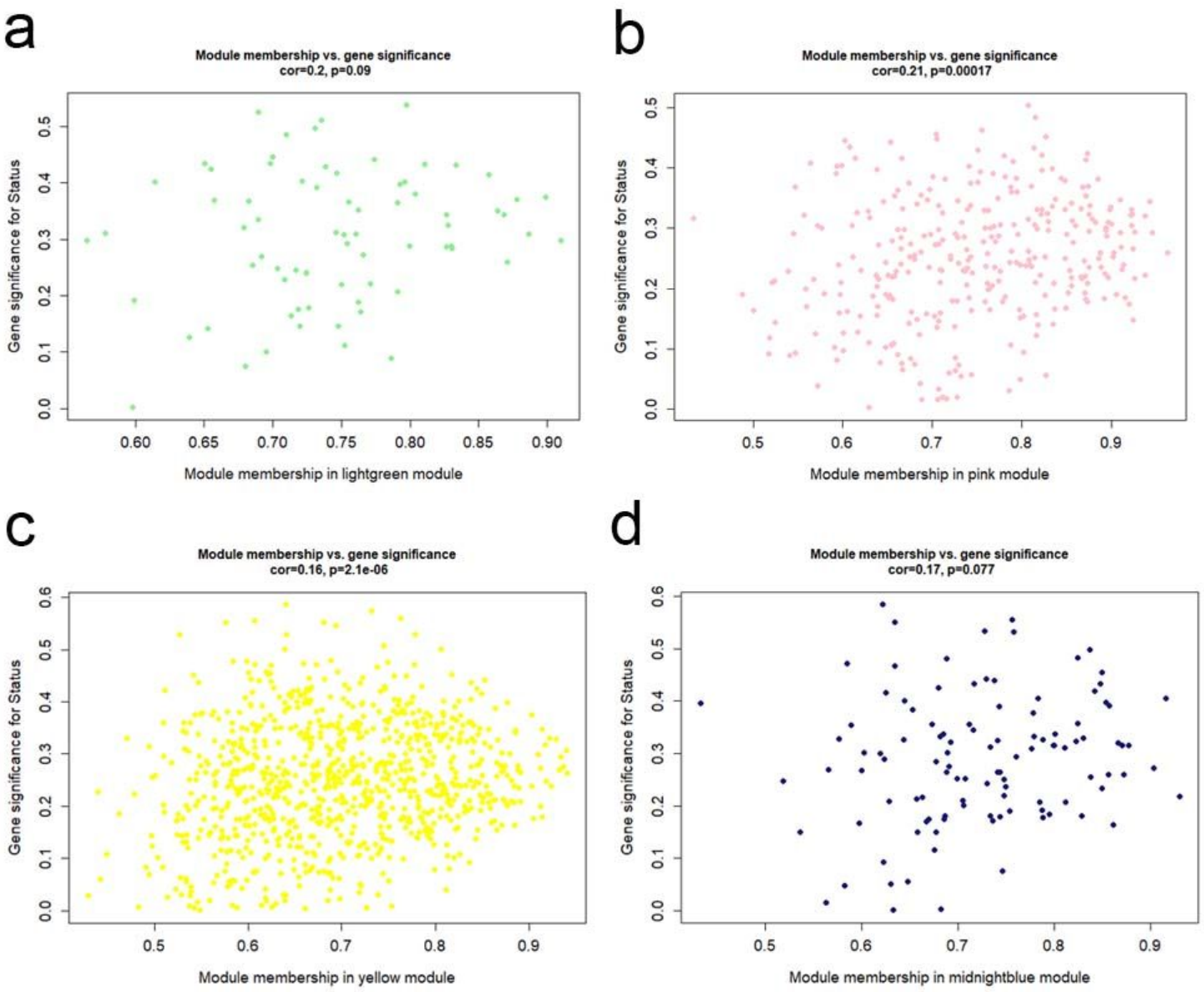

Figure 5

Module membership and gene significance. Scatterplot of Gene Significance (GS) for stroke vs. module membership (MM): a lightgreen module; b pink module; c yellow module; d midnightblue module

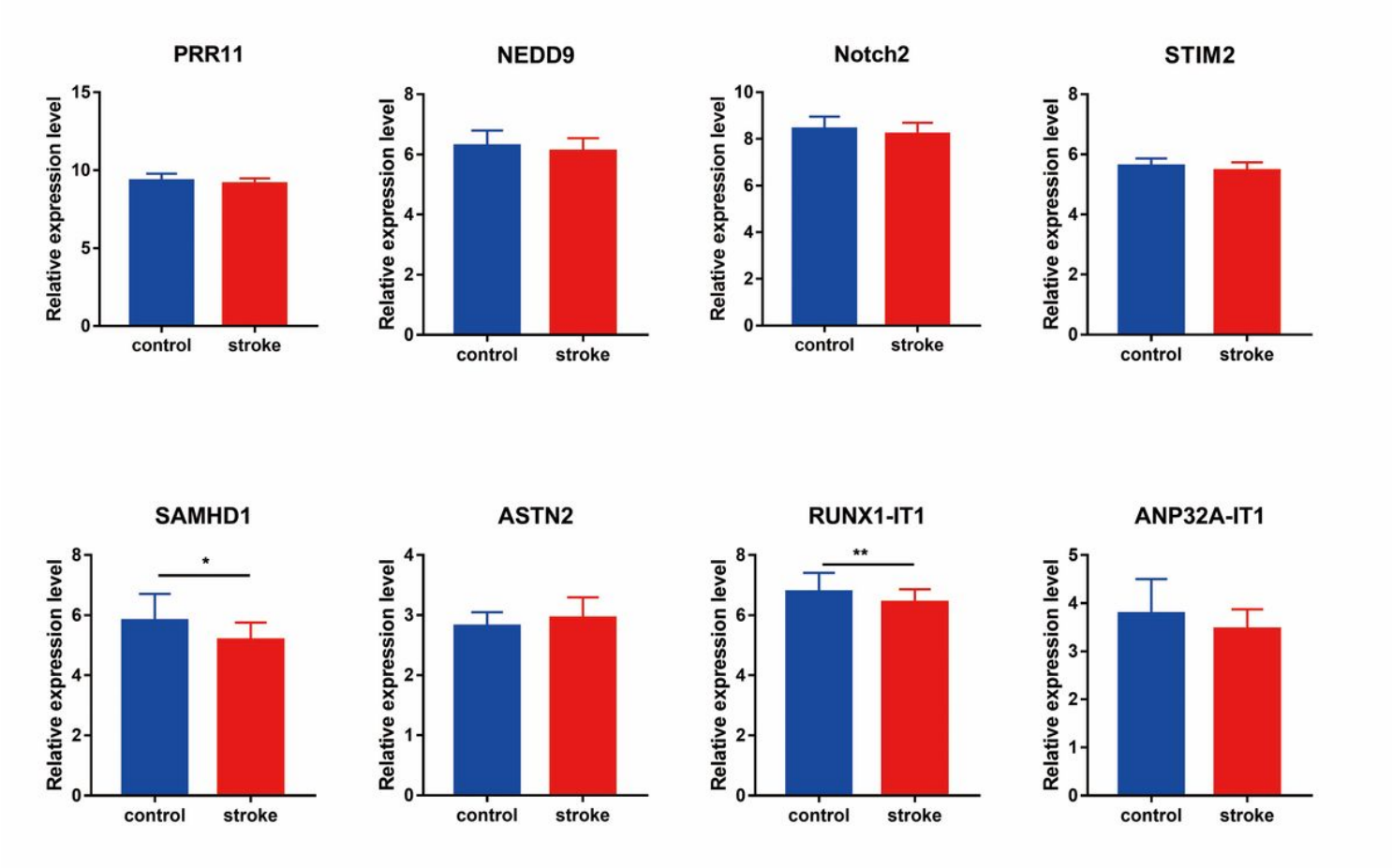

Figure 6

Hub genes expression in GSE22255. The barplot of the eight hub genes expression was produced. Among them, the expression levels of SAMHD1 and RUNX1-IT1 were significantly lower in stroke samples compared with control samples of GSE22255 dataset

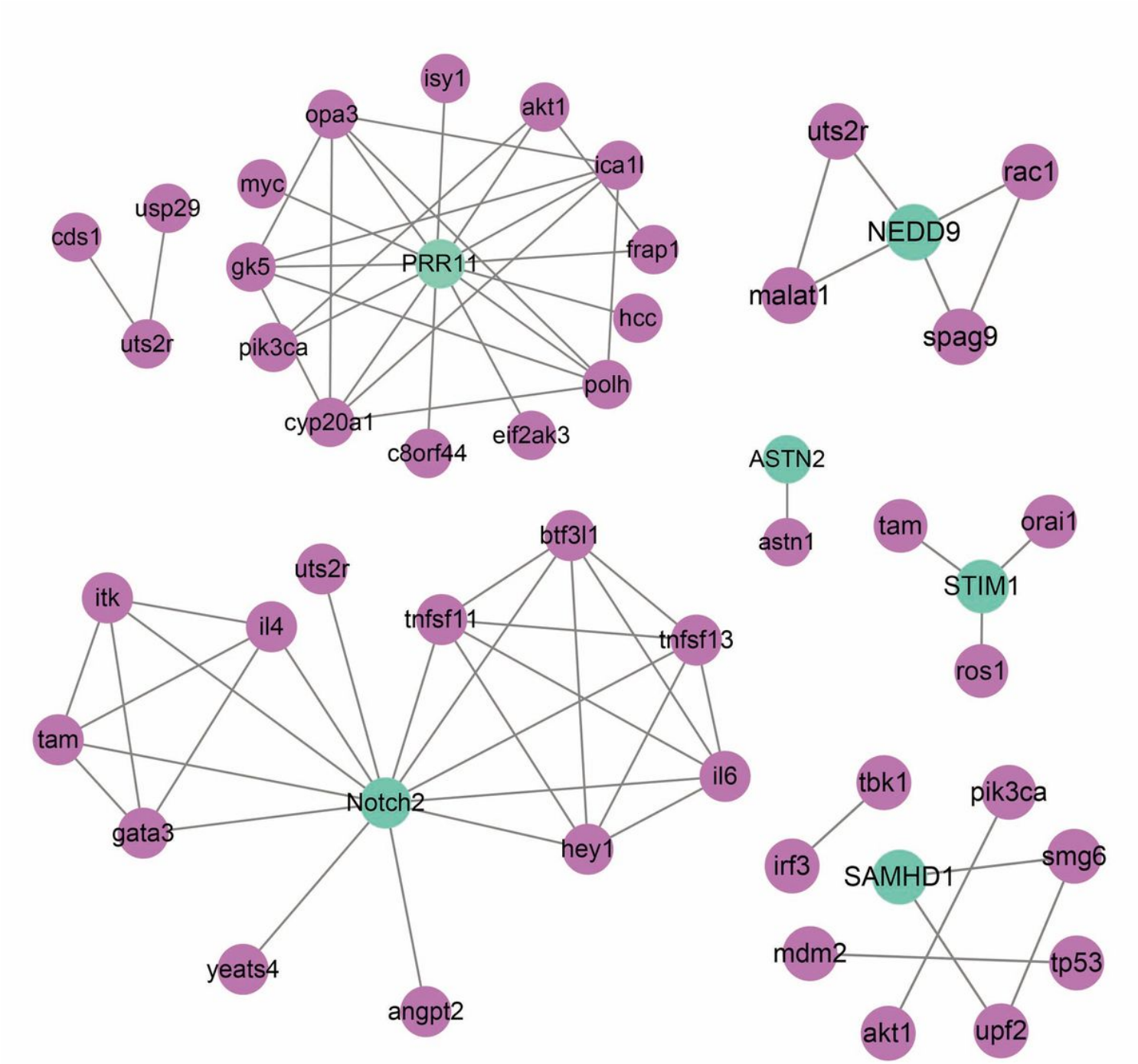

Figure 7

Validation for Hub genes regulatory network. Eight hub genes in the pink and yellow modules were colored with lightcyan, and the target genes were colored with magenta

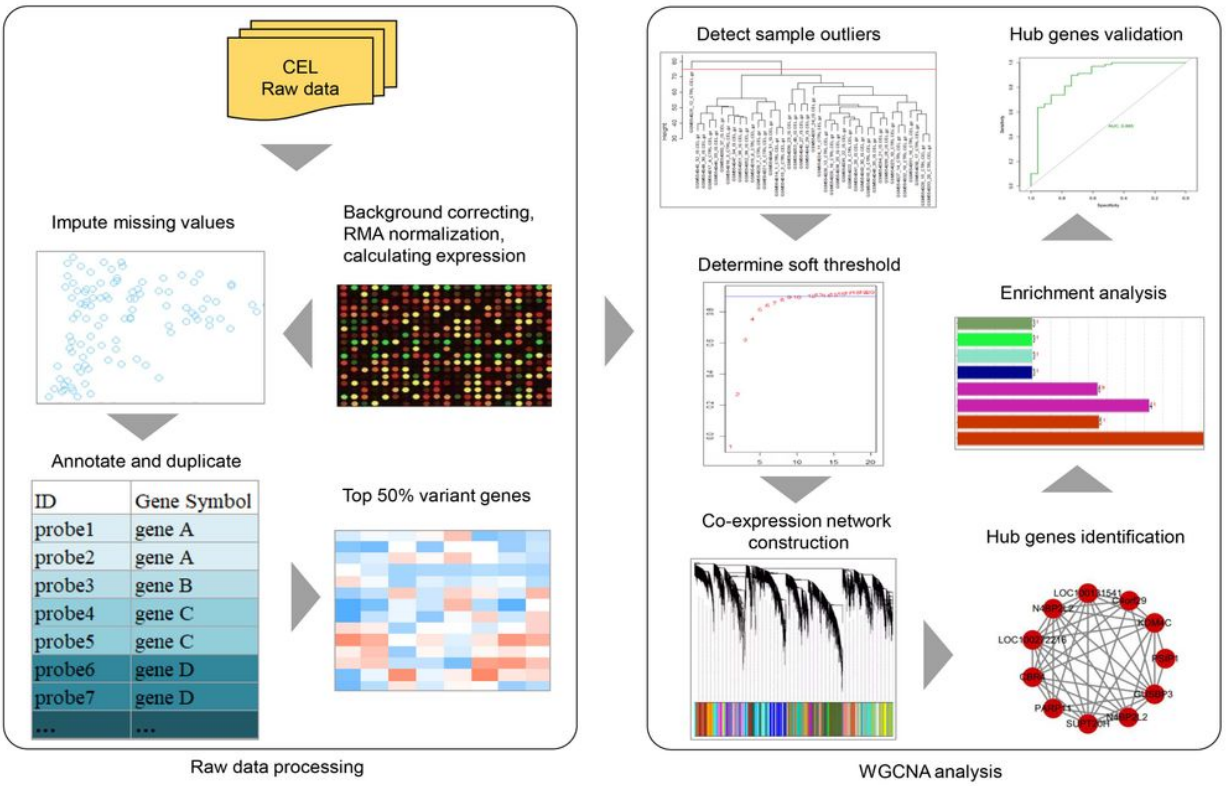

Figure 8

Workflow of raw data processing, WGCNA analysis and validation

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- ESM1.pdf.docx