Weighted gene co-expression network analysis of modules and biomarkers in early-phase acute lung injury with time-series gene expression profile

Jing Li
The First Affiliated Hospital of Zhengzhou University
Yanming Yang
The First Affiliated Hospital of Henan University of Chinese Medicine
Lili Fan
Henan Normal University
Zhengjun Cui (czj0371456@163.com)
The First Affiliated Hospital of Zhengzhou University

Research Article

Keywords: Acute lung injury, WGCNA, time-series expression profile, circadian rhythm, cell death

Posted Date: May 9th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1300671/v2

Abstract

Background: The pathogenesis of acute lung injury (ALI) is affected by a complex network of molecular factors and/or pathways; however, the time-series molecular mechanisms profile underlying ALI are unclear. Therefore, Time-series gene expression dataset GSE2565 was obtained from Gene Expression Omnibus, and weighted gene co-expression network analysis (WGCNA) was performed to identify time-series hub modules and genes and explore their associations with the progression of ALI.

Methods: Time-series gene expression dataset GSE2565 was obtained from Gene Expression Omnibus, the data were normalized, differentially expressed genes were screened, and WGCNA was performed to identify hub modules. In addition, Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathway enrichment analyses were conducted, and integration of module analysis and CytoHubba analysis were used to identify hub genes. Finally, receiver operating characteristic (ROC) curve analysis was used to measure the predictive accuracy of the hub genes.

Results: Module–trait analysis revealed that the red module, which was negatively correlated with the 8-h analysis time-point, was mainly involved in the regulation of the circadian rhythm, whereas the pink module, which was positively correlated with the 8-h analysis time-point, was mainly involved in regulation of cell death. Five hub genes, *Bnip3*, *Cdh11*, *Fam134b*, *Sult1a1*, and *Zbtb16*, were identified following ROC curve analysis.

Conclusions: our analysis of time-series expression data identified co-expression modules and pathways that were correlated with early-phase ALI. Moreover, the hub genes identified could provide new insights into the molecular mechanisms of ALI.

Introduction

The lungs are a primary target of many internal and/or external environmental insults, including microbial infection, autoantibodies, toxic gasses, pollutants, and gastric acids, and acute lung injury (ALI) often manifests after exposure to such insults. ALI can also evolve into a more severe condition known as acute respiratory distress syndrome (ARDS), which results in severe dysregulation of function and lung damage [1,2]. Although treatment methods are under investigation, the overall mortality of ALI and ARDS remains remarkably high in, for example, the USA (38.5% and 41.1%, respectively) and Shanghai (ARDS mortality rate in patients aged >15 years: 70%) [3]. Hence, therapeutic strategies to inhibit disease progression are required, and the underlying molecular mechanisms of the disease need further investigation.

ALI progression comprises three phases, namely, exudative, proliferative, and fibrotic phases, which are a continuum rather than a strict chronology of phases. The exudative phase (days 1–7) plays a fundamental role in the initiation of ALI. Without recovery during the exudative phase, the condition in some patients may evolve to the fibrotic phase, which is characterized by fibrosis and other irreversible pathological changes. Thus, timely source control should be at the core of ALI treatment, and

understanding the time-related pathogenesis of ALI, which is linked to outcomes, is important for current and future treatments [3-5].

Previous studies have suggested that various biological processes and multiple molecular factors are involved in the progression of ALI/ARDS pathogenesis, e.g., inflammatory responses, oxidative stress, apoptosis, autophagy, CCN1, and HMGB1 [6-12]. Therefore, the pathogenesis of ALI/ARDS is seemingly affected by a complex network of factors rather than by a single molecular factor and/or pathway at a specific time point, indicating that research that provides understanding of the time-series molecular mechanisms underlying ALI/ARDS must be conducted [13]. The roles of potential biomarkers in the progression of ALI have been studied, and various genetic factors have been associated with susceptibility to the disease, including specific genes such as *MMP3*, *Timp1*, *Ly6i*, and *Cxcl1* [14,15]. However, limiting research to specific genes will not reveal the complete mechanism underlying the progression of ALI because some important molecular regulators or pathways associated with disease progression will be neglected. In the current study, we used the time-serise gene expression dataset GSE2565 [16], which includes an archive of the changes in early-stage ALI (hours 0.5–72) during its progression. To understand the role of gene clusters in ALI progression, this dataset was analyzed for patterns of change. The GSE2565 data is considered part of the “big data” domain and includes a time-related gene expression pattern; hence, it was necessary to use weighted gene coexpression network analysis (WGCNA), which is an analysis approach used to evaluate the correlation between gene clusters and traits based on a coexpression network. WGCNA provides systems-level insights through recognition of hub modules and confirmation of key genes in time-series microarrays; thus, it can be used to identify potential candidate biomarkers or worthwhile targets, which in turn improve the probability of rapid intervention and clinical outcomes [17-19].

In the present study, differentially expressed genes (DEGs) were identified from the gene expression dataset GSE2565 using maSigPro [20,21]. WGCNA was used to process the DEGs, and hub modules associated with the progression of ALI were identified. In addition, pathway analysis of the hub modules was conducted. Furthermore, hub genes were distinguished from hub modules using CytoHubba analysis via Cytoscape [22]. Finally, the predictive ability of the hub genes was determined, using receiver operating characteristic (ROC) curve analysis.

Methods

Acquisition of GEO datasets

Microarray datasets of ALI were screened from the NCBI database (http://www.ncbi.nlm.nih.gov/geo). The time-series gene expression dataset GSE2565 on the GPL339 platform (Affymetrix Mouse Expression 430A Array), which was submitted by Sciuito [16], was extracted. The ALI data associated with GSE2565 came from 48 phosgene-exposed samples and 56 air-exposed samples. To increase the robustness of our analysis, biological duplicates

were also selected. Thus, 48 paired biological samples from male CD-1 mice (sampled at 0.5-, 1-, 4-, 8-, 12-, 24-, 48-, and 72-h postexposure) were obtained and analyzed using WGCNA (see Additional file 1: Table S1).

Microarray data preprocessing

Series matrix files obtained from the GEO website were used for further analysis after $\log_2$ transformation. Probe sets were annotated with Ensemble gene IDs using the data tables of an annotation file downloaded from the GEO website. In R, the ComBat function of the SVA package[23,24] was used for time-batch normalization to remove the batch effects of gene expression data from the 48 samples sampled at the 8 aforementioned time-points .

DEG analysis of the time-series dataset

MaSigPro from Bioconductor, an R package for analysis of time-series microarray data, was used to identify DEGs across the eight time-points in mice with ALI [20,21]. Typically, maSigPro finds genes with significant expression changes using a two-step regression approach: (1) a least-squares technique is used to select statistically significant genes between the control group and any other experimental group and (2) the conditions under which genes show statistically significant expression changes are identified. However, only step 1 for gene selection was performed in the present study to ensure that the loss of information was minimized. To select significant genes, a four-element regression model was defined, and a false discovery rate ($Q < 0.05$) was applied.

Coexpression network construction

The genes selected using maSigPro were used to construct a weighted gene coexpression network via the "WGCNA" R package in R Studio (version 3.6.1). An adjacency matrix of gene similarity according to pairwise Pearson correlation analysis was produced, and an appropriate soft threshold of $\beta = 8$ was obtained using the "pickSoftThreshold" function to strengthen the matrix to a scale-free coexpression network. A topological overlap matrix (TOM) was transformed from the adjacency matrix, and genes were clustered into different modules that were detected using the dynamic tree cutting algorithm of WGCNA.

Identification of significant hub modules

Module eigengenes (MEs) calculated using principal component analysis represented the major component of each gene module. The correlation between MEs and clinical traits was estimated using module–trait relationship analysis, which enabled the identification of modules related to external traits. Gene significance (GS) was defined according to the correlation between the gene and

external trait. Module membership (MM) was determined according to the correlation between the expression of each gene and ME. Module significance (MS) was calculated as the average absolute GS of the genes within the module and used to identify significant modules with clinical traits.

Functional enrichment analysis of hub modules

To assess the biological significance of hub modules identified using WGCNA, the genes of the hub modules were mapped to Metascape (http://metascape.org) for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses. The cut-off values for terms were as follows: $P < 0.01$; minimum overlap: 3; minimum enrichment factor > 1.5. When more than 20 terms were identified in GO or KEGG analyses, the top 20 terms were chosen for visualization [25].

Module identification of hub genes with module analysis

The “exportNetworkToCytoscape” function was used to export a PPI network by edge and node list files generated from hub modules in a format suitable for importing to Cytoscape. Based on maximal clique centrality (MCC) scores, the 15 highest-scoring genes identified using the MCC algorithm of Cytoscape were considered candidate hub genes. For WGCNA, candidate hub genes were also determined using absolute MM ≥ 0.8 and absolute GS ≥ 0.2. Genes shared in the PPI network and in WGCNA were considered hub genes. Finally, to determine the predictive accuracy of the candidate genes, the area under the curve (AUC) of the ROC was calculated using the “pROC” package. Genes with AUC > 0.70 in ROC analysis were considered actual hub genes [26].

Results

DEGs in ALI

Following adjustment for batch effects and a standardization process, the gene expression values of the 48 samples from the GSE2565 dataset were determined, and boxplots showing results before and after normalization are shown in Figure 1. Using the maSigPro approach, 3005 DEGs were screened out for further analysis.

Weighted gene coexpression network

Microarray quality was evaluated using sample clustering. Outliers were not detected in the clusters, and all 48 tissue samples were suitable for hierarchical clustering tree (dendrogram) construction (Figure 2). The soft-power threshold $\beta = 8$ was set to ensure a scale-free network (Figure 3). Power estimation results are shown in Additional file 2: Table S2. The 3005 genes were clustered into 9 modules based on the TOM. Among the nine modules, genes in the gray module were not coexpressed; thus, they were excluded from subsequent analysis [27].

Module–trait relationships and hub modules

The modules associated with various time-points in the progression of ALI were confirmed using module–trait relationship analysis (Figure 4). The 8-h ($R = -0.58$, $P = 2e^{-5}$), 12-h ($R = -0.40$, $P = 0.005$), 48-h ($R = 0.37$, $P = 0.01$), and 72-h ($R = 0.41$, $P = 0.004$) time-points were all significantly correlated with the red module (Figure 4), and positive or negative correlations between the MEs and time-points were observed. In particular, the red module was significantly correlated with ALI at 8 h, and this trend gradually increased from 8 to 72 h (Figure 4). A gradually decreasing correlation of the pink module with the 8-h ($R = 0.51$, $P = 0.005$), 12-h ($R = 0.15$, $P = 0.3$), 48-h ($R = -0.34$, $P = 0.02$), and 72-h ($R = -0.49$, $P = 4e^{-4}$) time-points was observed. Notably, both the red and pink modules were more highly associated with the 8-h time-point than that of any other (Figure 5). Consequently, these were considered the hub modules linked to the development of ALI.

Pathway enrichment analysis of genes in the hub modules

The results of GO term and KEGG pathway analysis of the red module are shown in Figures 6 and 7, respectively. Annotation information of the genes in the hub modules is provided in Additional file: Tables S3 and S4. Analysis of the top three GO terms revealed that genes in the red module were mainly significantly enriched with the terms circadian regulation of gene expression, response to hormone, and regulation of cellular response to stress. In KEGG pathway analysis, circadian rhythm, propanoate metabolism, and the PPAR signaling pathway were significantly enriched (Figure 6b). In the pink module (Figure 7), 17 GO terms and 2 KEGG pathways were enriched, primarily including regulation of cytokine production, positive regulation of cell death, and glycerophospholipid metabolism.

Hub genes screened in module analysis

PPI networks of the genes in the blue and pink modules were respectively constructed based on WGCNA. Based on MCC scores, the 15 highest-scoring genes were selected as hub genes, and results for the blue and pink module are shown in Figure 8a and 8b, respectively. Using GS > 0.2 and MM > 0.8 in the coexpression network, genes selected from the modules were considered hub genes (Tables 1 and 2). According to these criteria, 11 hub genes, found in both the PPI network and coexpression network, were selected. As shown in Figure 9, ROC curve analysis was used to explore the predictions of candidate hub genes as ALI biomarkers. Five genes with AUC values >0.7, *Bnip3*, *Cdh11*, *Fam134b*, *Sult1a1*, and *Zbtb16*, had high predictive accuracy for the development of ALI and were considered actual hub genes.

Discussion

In the present study, 3005 significant genes potentially associated with ALI development were identified, WGCNA was used to identify modules that were significantly associated with ALI, and hub genes in ALI were confirmed using module analysis. As a methodology, maSigPro has advantages for analyzing gene expression changes over time [20]; thus, we used this method to select statistically significant genes. Especially, only the first step was performed for the present work when applying the maSigPro procedure in order to minimize the loss of gene information. In addition, WGCNA as an comprehensive method to identify coexpression modules based on the similarity of expression patterns of genes was utilized for identifying key modules and hub genes related to the development of ALI.

Unlike previous studies, in which a static or combined static time data were typically used, we made use of dynamic time-series gene expression data to establish the genes specifically affected by clinical traits and maximize the identification of genes that participate in ALI-related biological processes. Specifically, previous research has often involved using a static approach to explore the mechanisms and implications of temporally regulated gene expression in ALI, focusing on either a few genes associated with a certain time-point or gene datasets from specific time-frames in which data were integrated from other datasets that included various time-points. For instance, Chen et al[28] analyzed ALI-related genes by combining dataset GSE2411 containing data at 4-h post-intervention, GSE18341 at 2-h post-intervention, and GSE17355 at 0-, 1-, 4-, and 10-days post-intervention, and the three different datasets were considered as a whole. Most biological processes are dynamic and gene expression has a high degree of temporal and spatial specificity; thus, time-course gene expression data is most applicable to the analysis of gene expression patterns in disease states. Indeed, time-series gene expression data, compared with static gene expression data, are considered to contain more information on gene regulatory networks. Furthermore, static gene expression analysis may be unable to show temporal trends in gene expression, resulting in the loss of some coexpressed genes from results [29]. However, using time-series gene expression data, we were able to identify the key modules and genes representing the developmental process of ALI based on gene expression at eight time-points.

We identified two gene modules that were highly correlated with the development of ALI. The red and pink modules showed significant correlations in module–trait relationships analysis and had high MS; thus, they were treated as hub modules. The red and pink modules were negatively and positively correlated with the 8-h time-point, respectively. According to GO and KEGG enrichment analyses, genes in the red module were significantly associated with regulation of circadian rhythm. Previous studies have indicated that cellular and molecular circadian rhythms are elicited by inflammation during ALI, indicating circadian rhythm's important role in the progression of ALI [30,31]. Circadian rhythm-related molecules are associated with various biological functions and dysfunctions and have emerged as promising targets for the treatment of various biological disorders [32].

According to GO analysis, genes in the pink module were mainly related to the regulation of cytokine production and positive regulation of cell death. KEGG analysis revealed that genes in this module were enriched in glycerophospholipid metabolism and HTLV-I infection. Phosphatidic acid, one of the elements

of glycerophospholipid metabolism, is involved in the regulation of the mTOR pathway, which affects protein synthesis, autophagy, and mitochondrial metabolism [33,34]. Autophagy, a form of programmed cell death, plays dual roles in the pathophysiological process of ALI in terms of disease progression/inhibition according to different cell types, insults, etc. [35]. In future research, in-depth evaluation of the role of autophagy in ALI should be conducted.

In the current study, we identified five susceptibility genes for ALI: *Bnip3*, *Cdh11*, *Fam134b*, *Sult1a1*, and *Zbtb16*. Bnip3 protein can trigger cell death through the pathways of necrosis, apoptosis, and autophagy [36] and exerts a dual effect through autophagy and apoptosis. The homeostasis of mitochondria maintains cell survival through autophagy, whereas apoptosis is mediated by the interaction between Bnip3 and mitochondrial fusion protein optic atrophy 1 to induce cell death [37]. Notably, Bnip3 expression levels have been reported to significantly increase following ALI [38].

Cdh11 encodes a type II classical cadherin that participates in the regulation of calcium-dependent cell–cell adhesion. Cdh11 protein is associated with fibrosis, inflammation, cancer, and other pathological processes [39]. In a previous study[40], the overall proportion of CDH11+ cardiac mesenchymal cells was shown to increase from day 3 to day 7 and remain unchanged over time in a sham group, whereas a significant increase in Cdh11 expression was observed from as early as 3 days following injury in line with the time course of inflammation resolution, implying that cells associated with inflammatory infiltration, such as macrophages, neutrophils, and monocytes, result in Cdh11-mediated fibrotic remodeling in the heart . According to this study, the expression of Cdh11 is time-dependent, suggesting that targeting Cdh11 may allow timely intervention that helps curb inflammation and the progression of ALI to pulmonary fibrosis.

Since its discovery in mammalian cells, the molecule encoded by *Fam134b* has been widely recognized as a molecular receptor for endoplasmic reticulum (ER)-specific autophagy (ER-phagy). Fam134b activation may exert an important role in the regulation of ER-phagy as selective autophagy plays a dual function in cell survival. When cells are subjected to insults, such as oxidative stress, chemical stimulation, and calcium overload, ER-phagy may help the ER restore homeostasis or induce cell death according to the timing and/or magnitude of stimuli [41,42]. Melchiotti [43] reported that Fam134b is a potential molecular target in response to inflammation and is associated with cytokine secretion. As it is involved in a common pathway, Fam134b may be associated with ALI; however, the specific role played by Fam134b in ALI progression has yet to be elucidated.

Sulfotransferase 1a1 (Sult1a1), which is encoded by the drug-processing gene *Sult1a1*, is a phase II metabolic enzyme that functions in the metabolism and detoxification of multiple drugs and chemicals [44]. Guo [45] found that the expression and activity of Sult1a1 was manifested in circadian rhythmicity that was directly regulated by the clock protein Bmal1. Furthermore, a growing body of evidence suggests that drug-processing genes are largely expressed in circadian rhythms and rhythmic expressions, which are closely associated with the time-dependency of toxicity and tolerance

of drugs in the body [46]. Therefore, elucidating the molecular mechanisms of Sult1a1 in ALI could help determine the best timing for drug administration.

Zbtb16 (zinc finger and BTB domain containing 16), also known as promyelocytic leukemia zinc finger protein, was reported to be primarily expressed in the apical membrane of bronchioles and is involved in restraining Toll-like receptor-induced inflammation in hyperglycemic states; hence, Zbtb16 may serve as molecular marker in lung tissues [47,48].

The current study has some limitations. The microarray data used for analysis were extracted from open access databases and had a small sample size; thus, further time-course studies are needed to improve the robustness of our results. Although, the mechanisms of lung injury induced by various insults in animal models share common biological pathways, the current analysis was limited to animal models exposed to photogas. Thereby, additional time-series studies of ALI induced by other insults are required[49].

Conclusion

To the best of our knowledge, this was the first study in which the WGCNA method was used to identify the complete activated gene set during the progression of ALI based on time-series expression data. The circadian rhythm and autophagy were identified as essential components in the pathophysiological mechanisms underlying early-phase progression of ALI. Moreover, five hub genes (*Bnip3*, *Cdh11*, *Fam134b*, *Sult1a1*, and *Zbtb16*) were identified, further study of which will provide new insights into the molecular mechanisms of ALI.

Declarations

Availability of data and materials

The dataset analyzed in this study can be derived from public repositories: GSE2565(https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE2565). All data generated or analyzed during this study are available from the corresponding author on reasonable request.

Abbreviations

ALI: Acute lung injury; ARDS: Acute respiratory distress syndrome; DEGs: Differentially expressed genes; GO: Gene Ontology; GS: Gene significance ; KEGG: Kyoto Encyclopedia of Genes and Genomes; ROC: Receiver operating characteristic ; MCC: Maximal clique centrality; MEs: Module eigengenes; MM: Module membership; MS: Module significance ; PPI:Protein–protein interaction; TOM: Topological overlap matrix ; WGCNA: Weighted gene co-expression network analysis

Acknowledgements

We acknowledge GEO database for providing their platforms and contributors for uploading their meaningful datasets.

Funding

The present work was supported by grants from the National Natural Science Cooperation Foundation of China (grant no. U1604188)

Authors' contributions

LJ and YYM conceived the work. LJ wrote the manuscript. FLL, LJ, and YYM participated in data analysis and interpretation. FLL revised the manuscript and participated in discussion. CZJ supervised the study. All authors approved the submission of the manuscript.

Ethics approval and consent to participate

GEO belongs to public databases. The animals involved in the database have obtained ethical approval. Our study is based on open source data, so there are no ethical issues and other conflicts of interest.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Tables

Table 1 Results of GS and MM of hub genes in red module

Hub gene	GS.red	MM.red
Npas2	0.524992056	0.812244171
Bnip3	0.591542468	-0.829989803
Ncbp2	0.466903275	0.843884636
Ints6	0.560423345	-0.833301709
Tef	0.568211177	-0.829765767
Sult1a1	0.53098546	-0.854797912
Pnpla2	0.515505357	-0.855865699
Ucp2	0.538712765	-0.827160254
Gyg	0.54025136	0.913996132
Hp	0.395057092	0.815863357
Map3k6	0.420717642	-0.855985892
Cdh11	0.505636263	0.861182638
H6pd	0.416300045	-0.856911322
Pcbd2	0.59191669	0.812055206
Ckb	0.38803167	0.80548476
Ppl	0.420790992	-0.833715548

Genes marked in bold black are common genes shared in PPI network

Table 2 Results of GS and MM of hub genes in pink module

Hub gene	GS.pink	MM.pink
Bcar3	0.406937702	0.813736076
Zbtb16	0.432140632	0.812118149
Fam134b	0.298895263	0.852403597
Stx3	0.29487702	0.810492385
Mlxip	0.365950339	0.819157543
Klf9	0.576878737	0.929771399
Adrb2	0.362958439	0.868261175
Ctla2a	0.362446535	0.8291722
Map3k6	0.499957606	0.815306015
Chst15	0.296325821	0.837480532
Lpin2	0.505598078	0.876790573

Genes marked in bold black are common genes shared in PPI network

References

[1]Liu X, Zhang J, Xie W. The role of ferroptosis in acute lung injury. Mol Cell Biochem 2022;477:1453–61. https://doi.org/10.1007/s11010-021-04327-7.

[2]Gonçalves-de-Albuquerque CF, Silva AR, Burth P, Castro-Faria MV, Castro-Faria-Neto HC. Acute Respiratory Distress Syndrome: Role of Oleic Acid-Triggered Lung Injury and Inflammation. Mediators Inflamm 2015;2015:260465. https://doi.org/10.1155/2015/260465.

[3]He Y-Q, Zhou C-C, Yu L-Y, Wang L, Deng J-L, Tao Y-L, et al. Natural product derived phytochemicals in managing acute lung injury by multiple mechanisms. Pharmacol Res 2021;163:105224. https://doi.org/10.1016/j.phrs.2020.105224.

[4]Mokra D, Mikolka P, Kosutova P, Mokry J. Corticosteroids in Acute Lung Injury: The Dilemma Continues. Int J Mol Sci 2019;20:4765. https://doi.org/10.3390/ijms20194765.

[5]Mowery NT, Terzian WTH, Nelson AC. Acute lung injury. Curr Probl Surg 2020;57:100777. https://doi.org/10.1016/j.cpsurg.2020.100777.

[6]Matthay MA, Zemans RL, Zimmerman GA, Arabi YM, Beitler JR, Mercat A, et al. Acute respiratory distress syndrome. Nat Rev Dis Primers 2019;5:18. https://doi.org/10.1038/s41572-019-0069-0.

[7]Zhang H, et al. Reactive oxygen species stimulated pulmonary epithelial cells mediate the alveolar recruitment of FasL+ killer B cells in LPS-induced acute lung injuries. J Leukoc Biol 2018;104:1187–98. https://doi.org/10.1002/JLB.3A0218-075R.

[8]Nadeem A, et al. Glucose-6-phosphate dehydrogenase inhibition attenuates acute lung injury through reduction in NADPH oxidase-derived reactive oxygen species. Clin Exp Immunol 2018;191:279–87. https://doi.org/10.1111/cei.13097.

[9]Fan K, Lin L, Ai Q, Wan J, Dai J, Liu G, et al. Lipopolysaccharide-Induced Dephosphorylation of AMPK-Activated Protein Kinase Potentiates Inflammatory Injury via Repression of ULK1-Dependent Autophagy. Front Immunol 2018;9:1464. https://doi.org/10.3389/fimmu.2018.01464.

[10]Zhang D., et al. Autophagy maintains the integrity of endothelial barrier in LPS-induced lung injury. J. Cell. Physiol. 2018;233(1):688–698.

[11]Lee S, et al. Production and application of HMGB1 derived recombinant RAGE-antagonist peptide for anti-inflammatory therapy in acute lung injury. Eur J Pharm Sci. 2018 Mar 1; 114():275-284.

[12]Dong N, Ji D, Huang X, Ying Z, Wang X, Chen C. Lipopolysaccharide-induced CCN1 production enhances interleukin-6 secretion in bronchial epithelial cells. Cell Biol Toxicol. 2018 Feb;34(1):39-49

[13]Patel VJ, Biswas Roy S, Mehta HJ, Joo M, Sadikot RT. Alternative and Natural Therapies for Acute Lung Injury and Acute Respiratory Distress Syndrome. Biomed Res Int 2018;2018:2476824. https://doi.org/10.1155/2018/2476824.

[14]Artham S, Verma A, Newsome AS, Somanath PR. Patients with acute respiratory distress syndrome exhibit increased stromelysin1 activity in the blood samples. Cytokine 2020;131:155086. https://doi.org/10.1016/j.cyto.2020.155086.

[15]Mao K, Geng W, Liao Y, Luo P, Zhong H, Ma P, et al. Identification of robust genetic signatures associated with lipopolysaccharide-induced acute lung injury onset and astaxanthin therapeutic effects by integrative analysis of RNA sequencing data and GEO datasets. Aging (Albany NY) 2020;12:18716–40. https://doi.org/10.18632/aging.104042.

[16]Sciuto AM, Phillips CS, Orzolek LD, Hege AI, Moran TS, Dillman JF 3rd. Genomic analysis of murine pulmonary tissue following carbonyl chloride inhalation. Chem Res Toxicol 2005;18:1654–60. https://doi.org/10.1021/tx050126f.

[17]Wu Z, Hai E, Di Z, Ma R, Shang F, Wang Y, et al. Using WGCNA (weighted gene co-expression network analysis) to identify the hub genes of skin hair follicle development in fetus stage of Inner Mongolia cashmere goat. PLoS One 2020;15:e0243507. https://doi.org/10.1371/journal.pone.0243507.

[18]Liang J-W, Fang Z-Y, Huang Y, Liuyang Z-Y, Zhang X-L, Wang J-L, et al. Application of Weighted Gene Co-Expression Network Analysis to Explore the Key Genes in Alzheimer's Disease. J Alzheimers Dis 2018;65:1353–64. https://doi.org/10.3233/JAD-180400.

[19]Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics 2008;9:559. https://doi.org/10.1186/1471-2105-9-559.

[20]Conesa A, Nueda MJ, Ferrer A, Talón M. maSigPro: a method to identify significantly differential expression profiles in time-course microarray experiments. Bioinformatics 2006;22:1096–102. https://doi.org/10.1093/bioinformatics/btl056.

[21]Yin L, Wang Y, Guo X, Xu C, Yu G. Comparison of gene expression in liver regeneration and hepatocellular carcinoma formation. Cancer Manag Res 2018;10:5691–708. https://doi.org/10.2147/CMAR.S172945.

[22]Tong Y, Song Y, Xia C, Deng S. Theoretical and in silico Analyses Reveal MYC as a Dynamic Network Biomarker in Colon and Rectal Cancer. Front Genet 2020;11:555540. https://doi.org/10.3389/fgene.2020.555540.

[23]Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882-3.https://doi.org/10.1093/bioinformatics/bts034.

[24]Peng X-Y, Wang Y, Hu H, Zhang X-J, Li Q. Identification of the molecular subgroups in coronary artery disease by gene expression profiles. J Cell Physiol 2019:10.1002/jcp.28324. https://doi.org/10.1002/jcp.28324.

[25]Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat Commun 2019;10:1523. https://doi.org/10.1038/s41467-019-09234-6.

[26]Al-Mansour MR, Wu J, Gagnon G, Knee A, Romanelli JR, Seymour NE. Linear versus volumetric CT analysis in predicting tension-free fascial closure in abdominal wall reconstruction. Hernia 2021;25:91–8. https://doi.org/10.1007/s10029-020-02349-6.

[27] Li J, Zhou D, Qiu W, Shi Y, Yang J-J, Chen S, et al. Application of Weighted Gene Co-expression Network Analysis for Data from Paired Design. Sci Rep 2018;8:622. https://doi.org/10.1038/s41598-017-18705-z.

[28]Cao F, Wang C, Long D, Deng Y, Mao K, Zhong H. Network-Based Integrated Analysis of Transcriptomic Studies in Dissecting Gene Signatures for LPS-Induced Acute Lung Injury. Inflammation. 2021;44(6):2486-2498. doi:10.1007/s10753-021-01518-8

[29]Ganesh SK, Joo J, Skelding K, Mehta L, Zheng G, O'Neill K, et al. Time course analysis of gene expression identifies multiple genes with differential expression in patients with in-stent restenosis. BMC Med Genomics 2011;4:20. https://doi.org/10.1186/1755-8794-4-20.

[30]Haspel JA, Chettimada S, Shaik RS, Chu J-H, Raby BA, Cernadas M, et al. Circadian rhythm reprogramming during lung inflammation. Nat Commun 2014;5:4753. https://doi.org/10.1038/ncomms5753.

[31]Yu D, Fang X, Xu Y, Xiao H, Huang T, Zhang Y, et al. Rev-erbα can regulate the NF-κB/NALP3 pathway to modulate lipopolysaccharide-induced acute lung injury and inflammation. Int Immunopharmacol 2019;73:312–20.

https://doi.org/10.1016/j.intimp.2019.04.035.

[32]Huang S, Jiao X, Lu D, Pei X, Qi D, Li Z. Recent advances in modulators of circadian rhythms: an update and perspective. J Enzyme Inhib Med Chem. 2020;35(1):1267-86. https://doi:10.1080/14756366.2020.1772249.

[33]Lutkewitte AJ, Finck BN. Regulation of Signaling and Metabolism by Lipin-mediated Phosphatidic Acid Phosphohydrolase Activity. Biomolecules 2020;10:1386. https://doi.org/10.3390/biom10101386.

[34]Zeng C, Wen B, Hou G, Lei L, Mei Z, Jia X, et al. Lipidomics profiling reveals the role of glycerophospholipid metabolism in psoriasis. Gigascience 2017;6:1–11. https://doi.org/10.1093/gigascience/gix087.

[35]Wang K, Chen Y, Zhang P, Lin P, Xie N, Wu M. Protective Features of Autophagy in Pulmonary Infection and Inflammatory Diseases. Cells 2019;8:123. https://doi.org/10.3390/cells8020123.

[36] Vishnupriya S, Priya Dharshini LC, Sakthivel KM, Rasmi RR. Autophagy markers as mediators of lung injury-implication for therapeutic intervention. Life Sci 2020;260:118308. https://doi.org/10.1016/j.lfs.2020.118308.

[37]Liu KE, Frazier WA. Phosphorylation of the BNIP3 C-Terminus Inhibits Mitochondrial Damage and Cell Death without Blocking Autophagy. PLoS One 2015;10:e0129667. https://doi.org/10.1371/journal.pone.0129667.

[38]Zhang W, Zhang J. Dexmedetomidine preconditioning protects against lung injury induced by ischemia-reperfusion through inhibition of autophagy. Exp Ther Med 2017;14:973–80. https://doi.org/10.3892/etm.2017.4623.

[39]Chang SK, Noss EH, Chen M, Gu Z, Townsend K, Grenha R, et al. Cadherin-11 regulates fibroblast inflammation. Proc Natl Acad Sci USA. 2011;108(20):8402–07. https://doi: 10.1073/pnas.1019437108.

[40]Schroer AK, Bersi MR, Clark CR, Zhang Q, Sanders LH, Hatzopoulos AK, et al. Cadherin-11 blockade reduces inflammation-driven fibrotic remodeling and improves outcomes after myocardial infarction. JCI Insight 2019;4:e131545. https://doi.org/10.1172/jci.insight.131545.

[41]Jiang X, Wang X, Ding X, Du M, Li B, Weng X, et al. FAM134B oligomerization drives endoplasmic reticulum membrane scission for ER-phagy. EMBO J 2020;39:e102608. https://doi.org/10.15252/embj.2019102608.

[42]Mo J, Chen J, Zhang B. Critical roles of FAM134B in ER-phagy and diseases. Cell Death Dis 2020;11:983. https://doi.org/10.1038/s41419-020-03195-1.

[43]Melchiotti R, Puan KJ, Andiappan AK, Poh TY, Starke M, Zhuang L, et al. Genetic analysis of an allergic rhinitis cohort reveals an intercellular epistasis between FAM134B and CD39. BMC Med Genet 2014;15:73. https://doi.org/10.1186/1471-2350-15-73.

[44]James MO and Ambadapadi S. Interactions of cytosolic sulfotransferases with xenobiotics. Drug Metab Rev . 2013;45:401–14.

[45]Guo L, Yu F, Zhang T, Wu B. The Clock Protein Bmal1 Regulates Circadian Expression and Activity of Sulfotransferase 1a1 in Mice. Drug Metab Dispos 2018;46:1403–10. https://doi.org/10.1124/dmd.118.082503..

[46]Sukumaran S, Almon RR, DuBois DC, Jusko WJ. Circadian rhythms in gene expression: Relationship to physiology, disease, drug disposition and drug action. Adv Drug Deliv Rev 2010;62:904–17. https://doi.org/10.1016/j.addr.2010.05.009.

[47]Kim JH, Rasaei R, Park S, Kim JY, Na S, Hong SH. Altered Gene Expression Profiles in the Lungs of Streptozotocin-induced Diabetic Mice. Dev Reprod. 2020;24(3):197-205. https://doi:10.12717/DR.2020.24.3.197.

[48]Sadler AJ, Rossello FJ, Yu L, Deane JA, Yuan X, Wang D, et al. BTB-ZF transcriptional regulator PLZF modifies chromatin to restrain inflammatory signaling programs. Proc Natl Acad Sci U S A 2015;112:1535–40. https://doi.org/10.1073/pnas.1409728112.

[49] Mokra D, Mikolka P, Kosutova P, Mokry J. Corticosteroids in Acute Lung Injury: The Dilemma Continues. Int J Mol Sci 2019;20:4765. https://doi.org/10.3390/ijms20194765.

Figures

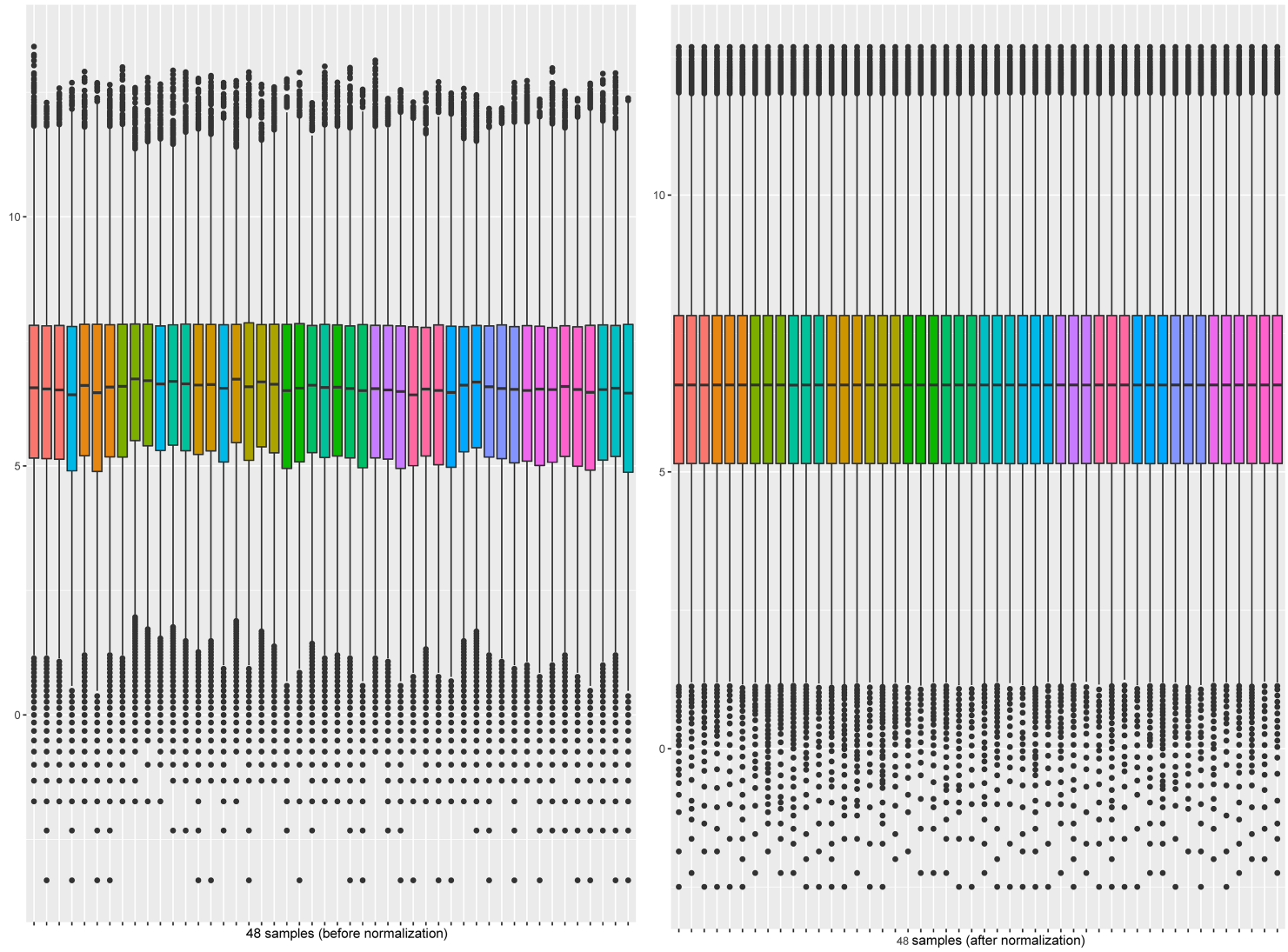

Figure 1

Box plot analysis for unnormalised and normalised data. Normalised gene expression levels of each data set are presents closer sequencing depth.

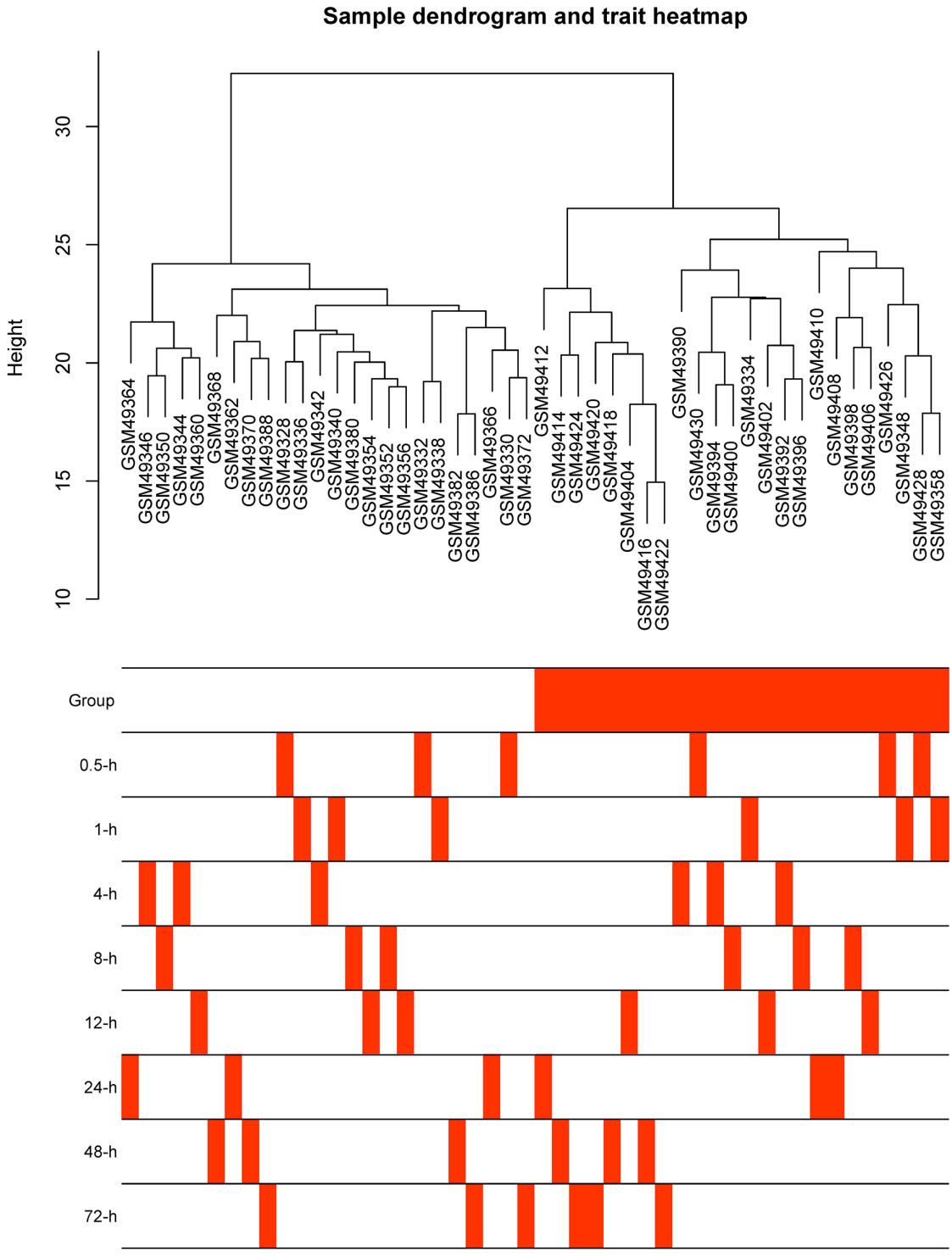

Figure 2

Sample dendrogram and trait heatmap. The clinical traits, the group and the different time points are presented at the bottom. The colors represent the proportion to clinical traits（group, time-points)

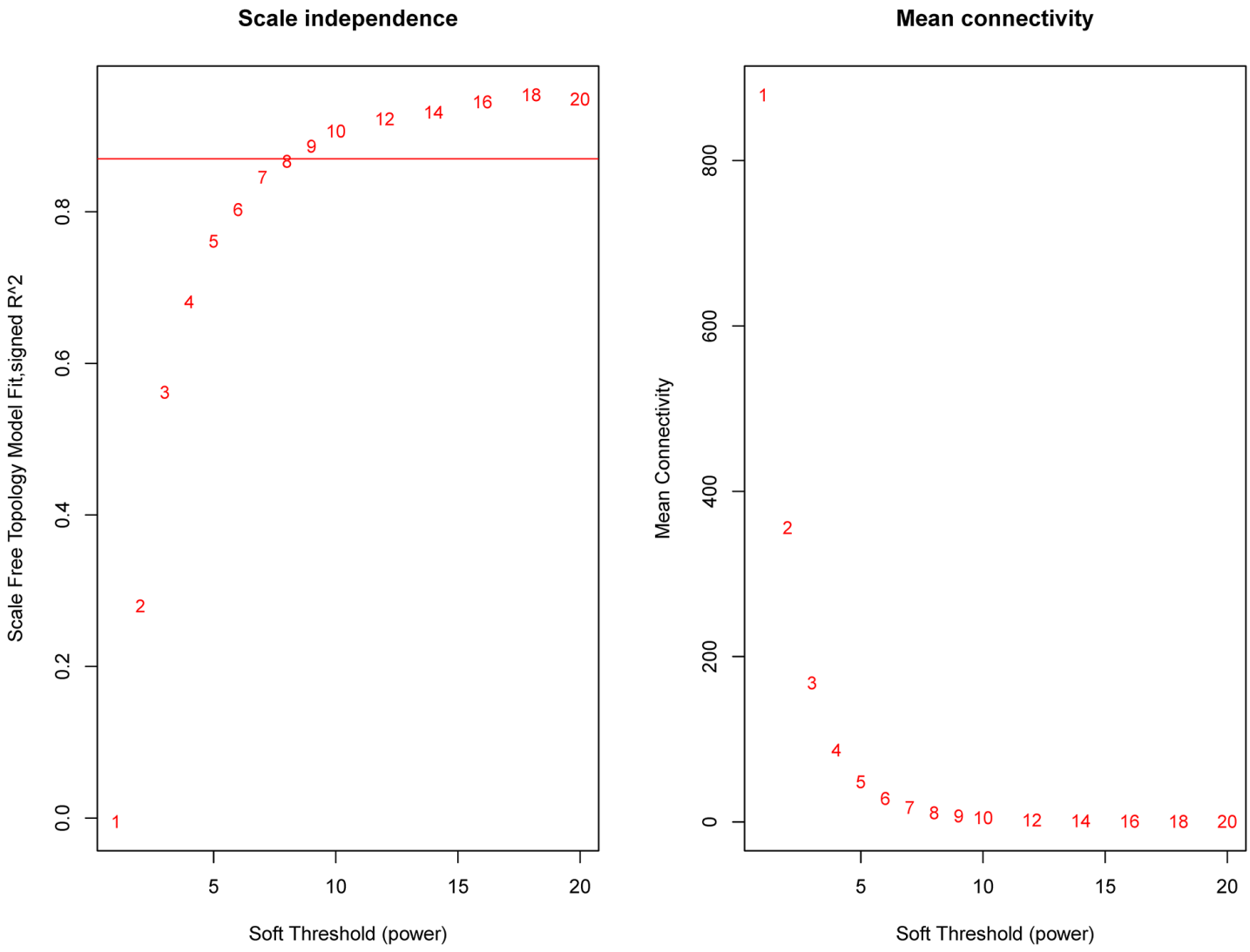

Figure 3

Selection for various soft-thresholding powers (β). The left panel shows the scale-free fit index, signed R2 values (y-axis) and the soft threshold power (x-axis).β = 8 was chosen for the subsequent analysis. The right panel shows that the mean connectivity (y-axis) is a strictly decreasing function of the power β (x-axis).

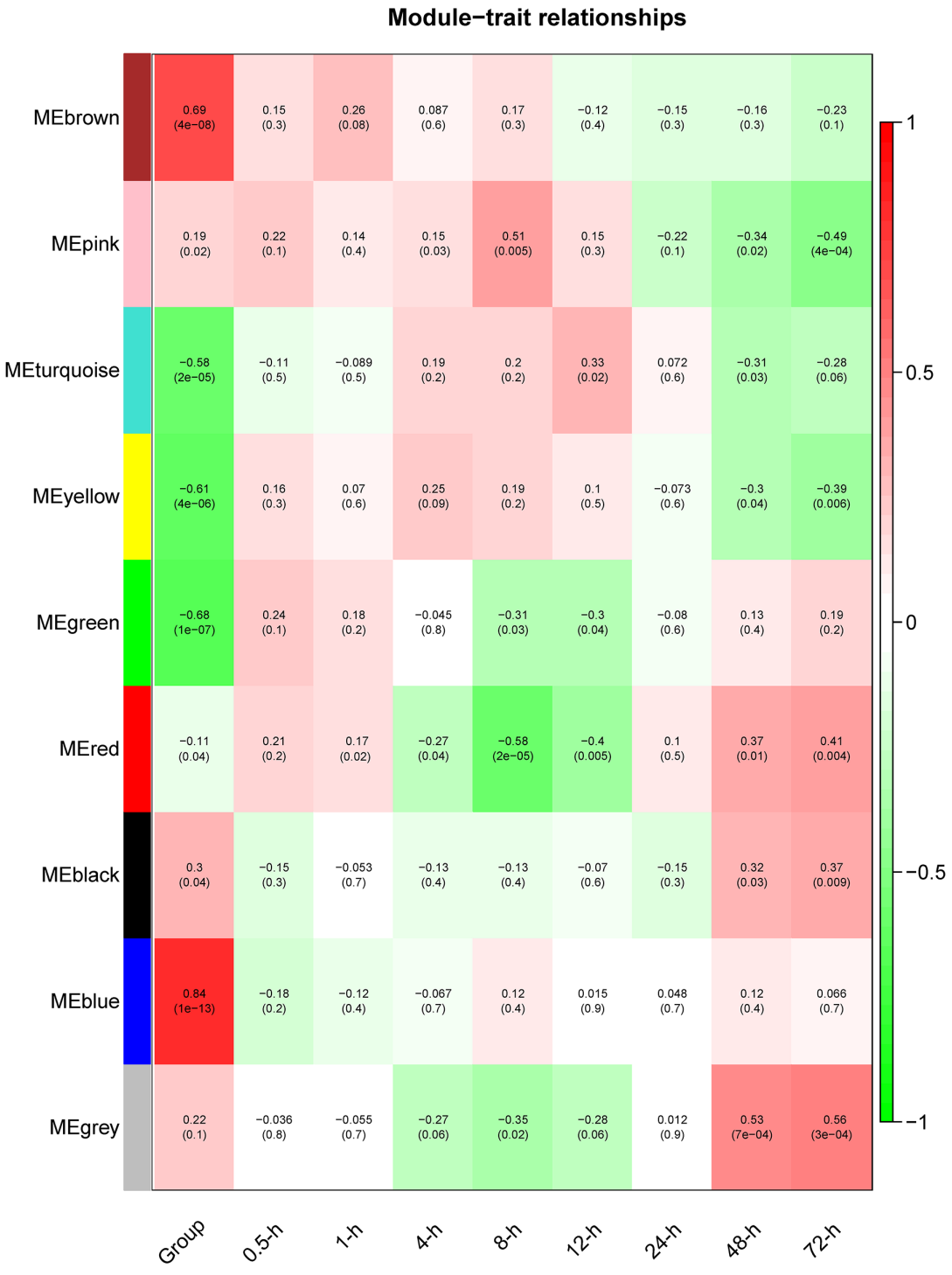

Figure 4

Heatmap of the correlation between module eigengenes and clinical traits. Each row corresponds to a module, and each column corresponds to a trait. Each cell contains the corresponding correlation and P value. The table is color-coded by correlation according to the color legend;green represents negative correlation and red represents positive correlation

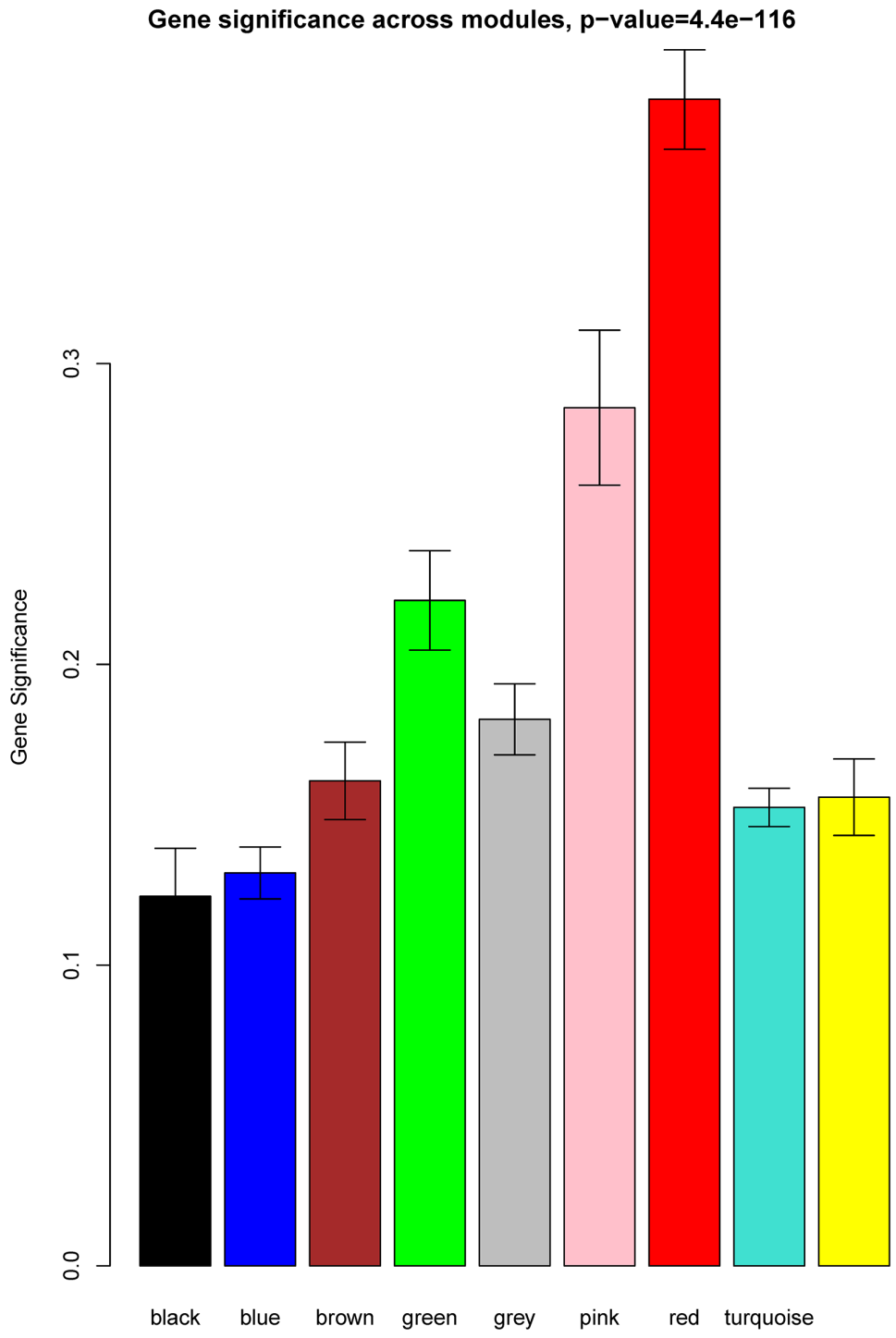

Figure 5

Distribution of average gene significance and errors in the modules associated with ALI.

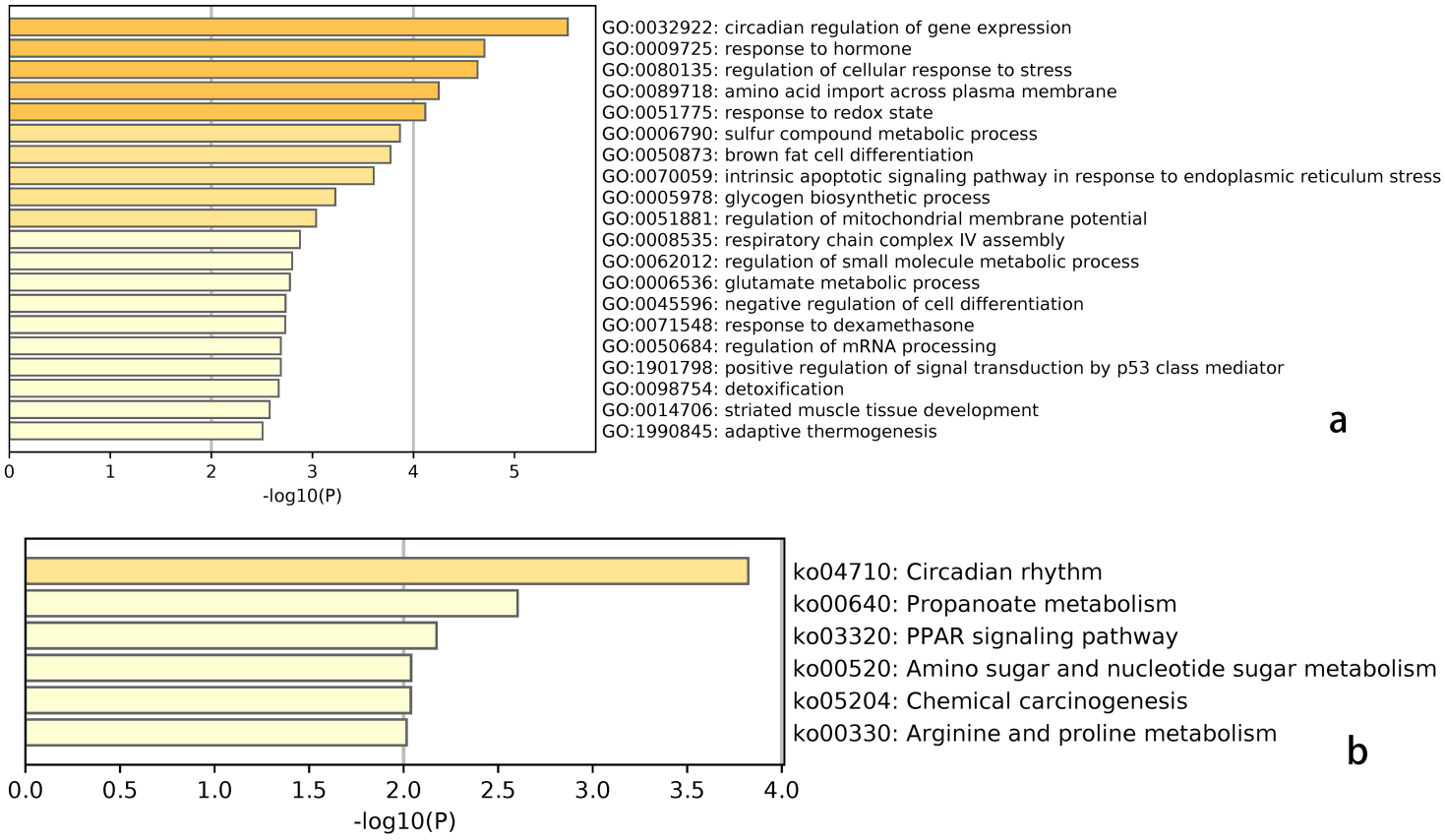

Figure 6

Functional enrichment analysis of the red module genes. (a) Gene Ontology analysis of genes in the red module. (b) KEGG pathway enrichment analysis of genes in the red module.

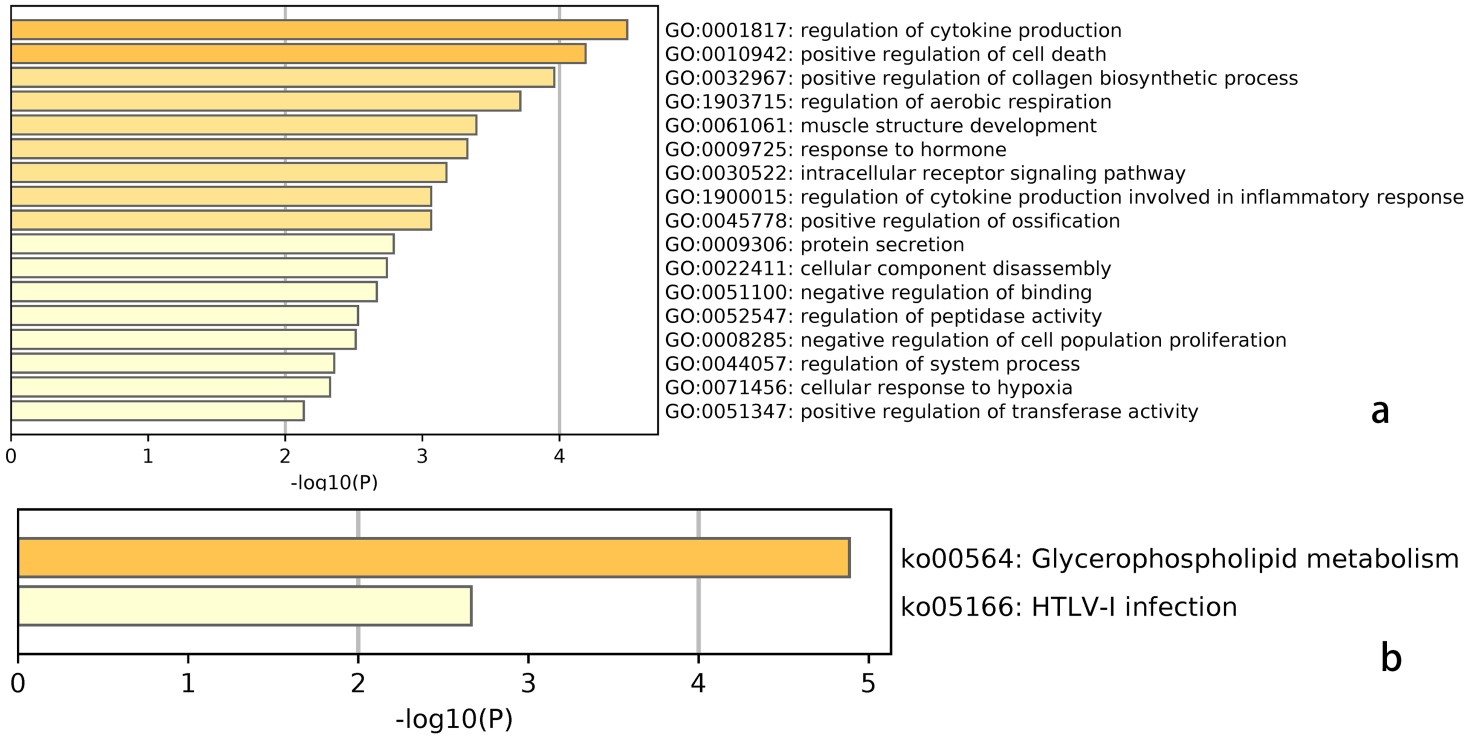

Figure 7

Functional enrichment analysis of the pink module genes. (a) Gene Ontology analysis of genes in the pink module. (b) KEGG pathway enrichment analysis of genes in the pink module.

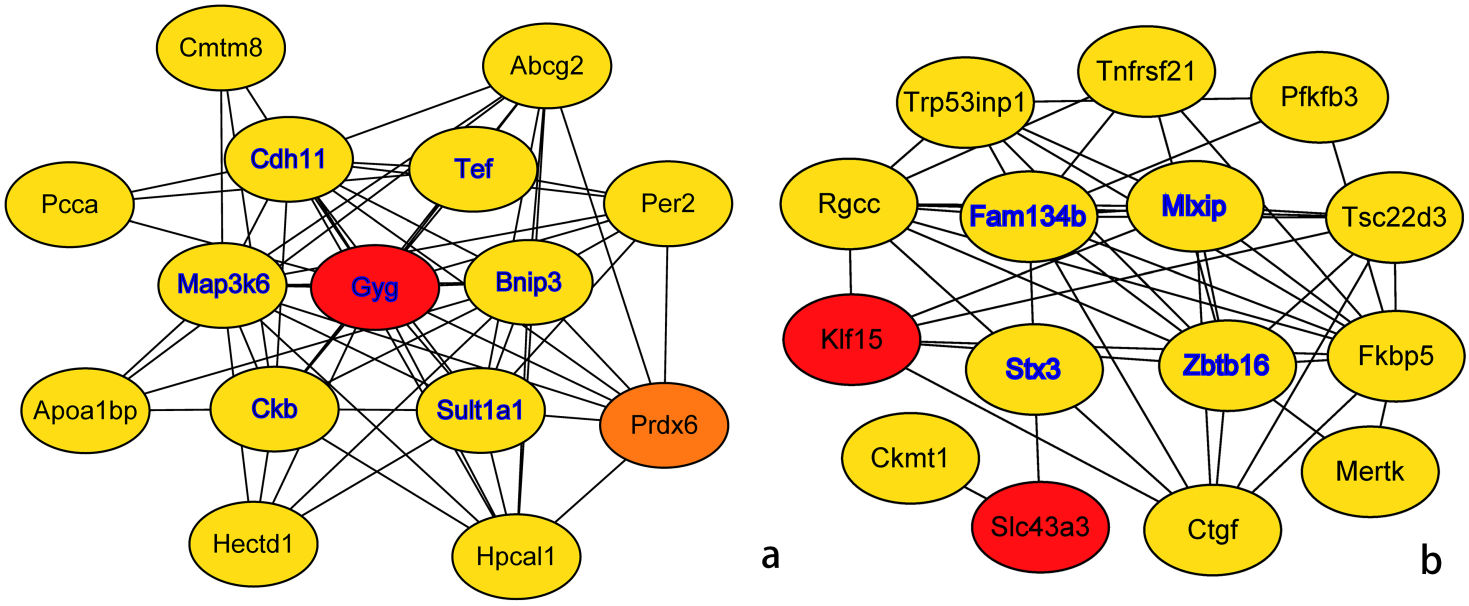

Figure 8

Identification of the hub genes from the PPI network using MCC algorithm.. The top 15 hub genes were visualized on the Cytoscape application. (a) Candidate hub genes in the red module. (b) Candidate hub genes in the pink module. Common hub genes shared in PPI network and co-expression network are marked in blue color.

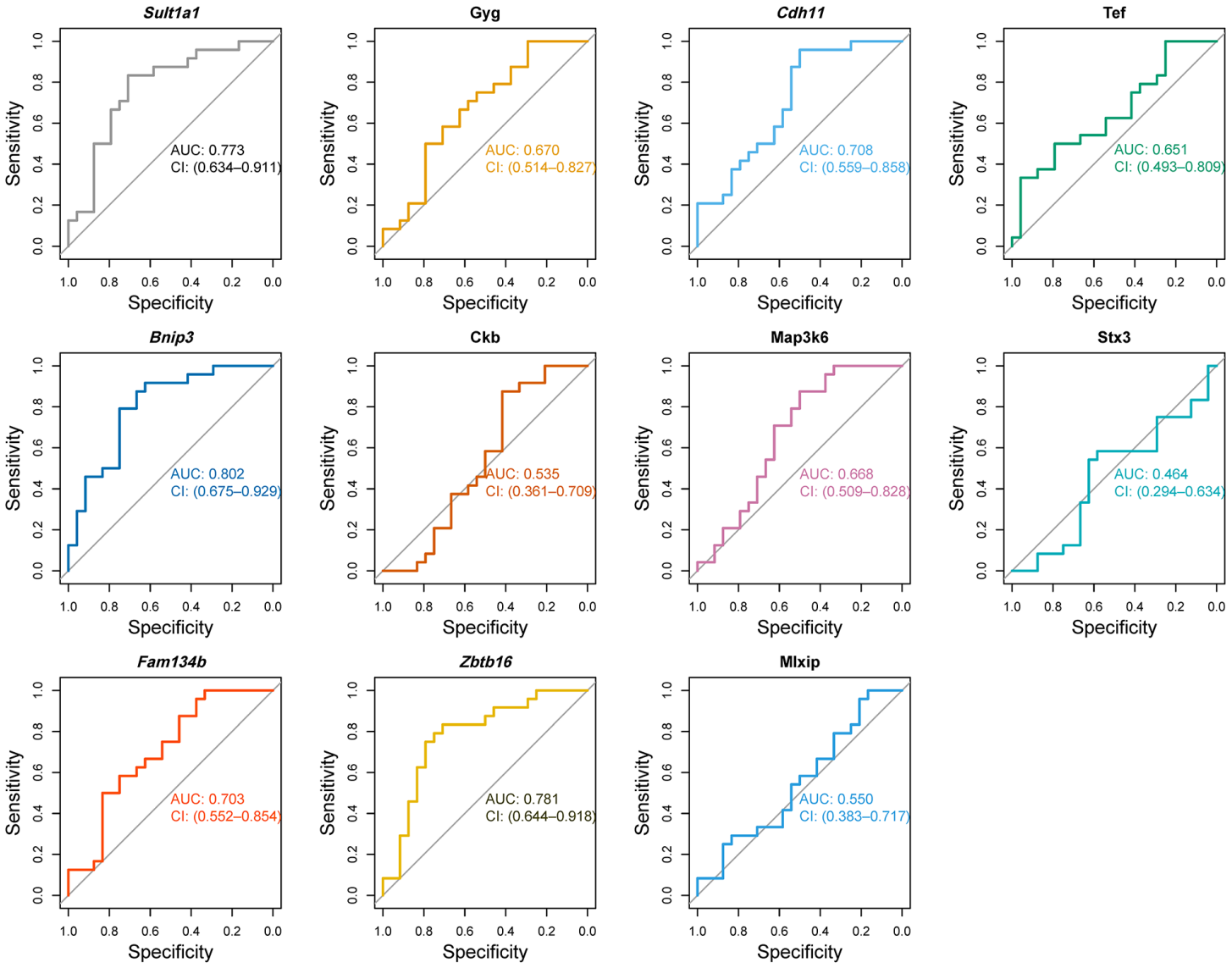

Figure 9

ROC analysis of hub genes. AUC > 0.7 indicated that the model had a good fitting effect

Supplementary Files

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

- tableS1InformationOfSamples.xlsx