Key Genes and Pathways in Human Left Atrial Appendage Identified by Co-expression Analysis

Yi-chi Yu
Department of Cardiology
Jian Sun
Department of Cardiology,
Ya-Qin Han
Department of Cardiology
Mu Chen
Department of Cardiology
Peng-Pai Zhang
Department of Cardiology
Quan-Fu Xu
Department of Cardiology
Rui Zhang
Department of Cardiology
Yi Yu
Department of Cardiology
Wei Li
Department of Cardiology
Qun-Shan Wang
Department of Cardiology
Mei Yang (yangmei@xinhuamed.com.cn)
Department of Cardiology
Yi-Gang Li (liyigang@xinhuamed.com.cn)
Department of Cardiology

Research

Keywords: left atrial appendage, LAA closure, WGCNA analysis, atrial fibrillation, neurohormonal changes

Posted Date: February 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-169544/v1

Abstract

Background and Objective

Left atrial appendage (LAA) closure (LAAC) is a technique that has shown potential for the prevention of thrombo-embolic events in atrial fibrillation (AF) patients. The short and long-term effects of LAAC on neuro-hormonal changes have been highlighted in several recent studies. In addition, metabolic and hemodynamic profile changes were identified after LAAC, which may be attributed to the potential influence of specific LAA genetic profiling. However, only a few studies have deciphered the specific LAA genetic profiling. Therefore, we sought to conduct a weighted gene co-expression network analysis (WGCNA) to identify highly correlated gene modules in the LAA and to identify the hub genes with the highest degree of connectivity in selected modules. Functional enrichment analysis was performed to investigate the pivotal biological processes and pathways of defined gene modules in the LAA.

Material and Methods

Genes exhibiting the highest expression levels (top 25%) of variation in the microarray samples from the combined GSE41177 and GSE79768 dataset were identified. These datasets were obtained from the Gene Expression Omnibus (GEO) database. The combined dataset, which was used to conduct the WGCNA, included 38 paired samples that compared LAA (n = 19) with left atrium-pulmonary vein junction (LA-PV, n = 19) specimens. Gene ontology and functional enrichment analyses were performed to define genes belonging to the key modules of LAA. Hub genes were screened out from the key modules by algorithms and interactions analysis and which were visualized using Cytoscape software.

Results

Two modules with 397 (pink) and 419 (black) probes were identified to be specifically related to LAA (pink: r = 0.22, p = 9.7×10 − 6; black: r = 0.27, p = 2×10 − 8). In the functional analyses, the pink module showed an association with serine-type endopeptidase activity, cell-cell adhesion, synapse and axon guidance and protein processing. The black module was primarily associated with metabolic processes, such as the triglyceride metabolic process, triglyceride biosynthetic process, fatty acid oxidation, carbohydrate biosynthetic process, insulin signaling pathway, regulation of lipid storage, PPAR signaling pathway, regulation of lipolysis adipocytes, response to peptide hormone and amino acid biosynthesis. A total of five genes, including LRRN4 (leucine-rich repeat neuronal protein 4) and KLK11 (kallikrein related peptidase 11) in the pink module, as well as GYG2 (glycogenin 2), GPD1 (glycerol-3-phosphate dehydrogenase 1) and DGAT2 (diacylglycerol O-acyltransferase 2) from the black module, were identified as hub genes.

Conclusions

Using WGCNA bioinformatic approach, we defined key genes and pathways with specific biological characteristics in the human LAA, and our results thus are helpful to understand the underlying mechanisms responsible for the neuro-hormonal changes following LAA closure procedures.

Introduction

Atrial fibrillation (AF) is one of the most significant causes of thromboembolic stroke in adults. The left atrial appendage (LAA) is a finger-like extension originating from the main body of the left atrium, and is recognized as the most important source of thrombus due to low levels of contractions and local blood stagnation in case of AF. Following left atrial appendectomy and epicardial LAA ligation strategies, LAA closure with endocardial LAA occluder devices, such as Watchman (Boston Scientific) and Lambre (LifeTech Scientific Co. Ltd.), has achieved equal benefits and is linked with fewer complications in the field of embolism prophylaxis.

It is known that intracardiac device implantation procedures like cardiac resynchronization therapy (CRT), are linked with post procedural changes on hemodynamic, metabolic and endocrine profiles (ref), in fact, these changes were also observed among patients post LAAC.[1] Intriguingly, in “LAA HOMEOSTASIS” study,[2] a prospective observational study, which compared the differences between patients with epicardial and endocardial LAAC, revealed more apparent neurohormonal changes in patients undergoing epicardial LAAC, especially on the renin-angiotensin-aldosterone system (RAAS), lipid and glucose metabolism, systemic blood pressure and natriuretic peptide levels. Since the difference between the two approaches is directly based on whether LAA is conserved or excluded, the physiological role of LAA on this process is highlighted.

Previous studies have attempted to reveal LAA specificity from the standpoint of genetic profiling. It is reported that the Pitx2 (paired-like homeodomain transcription factor 2) gene is highly enriched in the left atrium and might serve as a key regulator of left-right atrial asymmetry. In contrast, expression of Bmp10 (bone morphogenetic protein 10), a member of the TGFβ family, was significantly lower in left atria than in right atria.[3] In addition, expression of numerous genes, such as MYOZ1, C9orf3 and FANCC,[4] varied significantly in LAA, and might play significant role in the pathogenesis of AF. However, differential expression analysis allows each gene to be analyzed independently, but the interactions between genes remain elusive. It is well known that biological processes are dependent of numerous interactions between many cellular components and genes, a gene network analysis based on bioinformatics approaches could further provide valuable information on LAA in pathogenesis of various diseases.[5]

Weighted gene correlation network analysis (WGCNA) is a bioinformatics algorithm used to identify disease-related gene modules and the synergistic effect between key genes that contribute to phenotypic traits.[6,7] Instead of the traditional methods, which explore differentially expressed genes (DEGs) on individual gene, WGCNA can further construct scale-free networks and cluster highly correlated genes into different modules. In addition, WGCNA allows us to measure the relationships between modules (eigengene networks) and intramodular gene rankings determined through connectivity. In terms of

application, WGCNA has not been limited to genetic profiling and biomarker exploration in a range of diseases, notably cancers,[8–12] but has extended into mass spectrometry (MS)[13] and competing endogenous RNA (ceRNA)[14] research studies.

In this study, we performed a WGCNA analysis to identify relevant genes and potential pathways highly active in LAA, to determine the potential impact of LAA exclusion on systemic metabolism.

Materials And Methods

Dataset collection and pre-processing

Using the keywords “left atrial appendage” (Title) AND (“Series” (Entry type) AND “Homo sapiens” (Organism)), two microarray datasets containing the expression profiles of human LAA tissue were downloaded from the website of the Gene Expression Omnibus (GEO) database (http://www.ncbi.nlm.nih.gov/gds). These datasets were subsequently transformed into an expression matrix. The microarray dataset, GSE41177 (submitted by Yeh et al.)[16] consisted of 38 paired samples that compared samples from the LAA (n = 19) and left atrium-pulmonary vein junction (LA-PV, n = 19), while the other microarray dataset, GSE79768 (submitted by Tsai et al)[17] consisted of 26 paired samples from the LAA (n = 13) and right atrial appendage (RAA, n = 13). Both datasets were based on the GPL570 platform (Affymetrix Human Genome U133 Plus 2.0 Array), which contains a total of 54,675 probes. AF status, age and gender were extracted from the metadata for each sample. The two gene expression matrix files were merged into one file using the “sva” package in R (Version 3.6.1), which is used to eliminate batch effects and has been used in previous studies.[18,19] All batch normalized samples were further quartile normalized using preprocessCore tool package and are displayed in box plots (Figure S1). Gene ID conversion and quartile normalization were performed using the hgu133plus2.db database package. Genes were filtered out using an algorithm for the construction of the WGCNA. Additionally, the flashClust tool package in R was used with the appropriate threshold for sample clustering to detect and subsequently eliminate outliers.

Co-expression network construction through WGCNA

Power values were screened out in the construction of the co-expression modules using the WGCNA algorithm. Genes with the top 25% of variance were filtered using the algorithm and used for further analysis. A gradient test (power value from 1 to 20) was used for the analysis of scale independence and average connectivity of the modules with different power values. The appropriate power value was determined when the scale independence value was 0.9. Then, the weighted adjacency matrix was created using the formula $a_{mn} = |c_{mn}|\beta$(c_{mn}: Pearson's correlation between genes, a_{mn}: adjacency between genes, β: soft-power threshold). Then, the weighted adjacency matrix was transformed into a topological overlap measure (TOM) matrix to estimate its connectivity in the network. Average linkage hierarchical clustering was used to construct a clustering dendrogram of the TOM matrix. The minimal gene module size was set to 30, while the threshold for the merging of similar modules was set to 0.25.

Construction of weighted co-expression gene modules and identification of modules associated with clinical traits

Gene co-expression network analysis was performed using the WGCNA R software package. The optimal soft threshold for adjacency computation was graphically determined. The transformed expression matrix was entered into the WGCNA package and functions, modules, and corresponding Eigengenes were obtained. The cutree Dynamic function was used to cut off the branches in the gene hierarchical clustering dendrograms, which resulted in co-expression modules. The correlated modules ($r > 0.75$) were then merged. The dissimilarity of the module Eigengenes (ME) was calculated using the module Eigengenes function in the WGCNA R package. The association between Eigengene values with clinical traits was assessed using Pearson's correlation and then illustrated in heatmaps. All the modules that were significantly associated with LAA were selected for further analysis.

Hub gene mining and module visualization

The hub gene candidates of each LAA-associated module were determined using the following parameters calculated through the WGCNA package: module membership (MM), gene significance (GS) and intramodular connectivity (kME). In this study, the hub genes in the corresponding modules were firstly filtered out using a threshold of GS > 0.2 and MM > 0.8, and their importance was then determined by sorting using kME. The top 30 genes in each gene module of interest, which represented a potentially central status in the protein-protein interaction (PPI) network, were selected for further analysis and visualized using Cytoscape v3.7.0 (http://cytoscape.org).[20] The highest ranked nodes in the key modules were screened out as crucial genes (I) using three methods: the "chooseTopHubInEachModule" function in the WGCNA package, the highest intramodular connectivity and the "cytoHubba" plugin in Cytoscape.

Analysis of differentially expressed genes (DEGs)

The R package, "limma", was used to screen out differentially expressed genes (DEGs) from the dataset merged using the "sva" package for three sets of comparisons: LAA vs RAA, LAA vs LA-PV and RAA vs LA-PV. The genes meeting the cutoff criteria ($P < 0.05$ and |log fold change (logFC) |>1) were designated as DEGs. The DEGs were visualized as a volcano plot using the R package, "ggplot2". Subsequently, common DEGs that were involved in LAA and that overlapped with hub genes in the modules of interest, were also identified as potential "crucial genes"(II). The results were displayed as a Venn diagram using the R package, "VennDiagram".

Functional enrichment analysis of key modules and crucial genes

To understand the functional significance of the key modules and crucial genes (I and II from previous steps), GO (Gene Ontology) and KEGG (Kyoto Encyclopaedia of Genes and Genomes) functional enrichment analyses were performed on the LAA-associated modules using the clusterProfiler R package and org.Hs.eg.db database package in R language. The cutoff value for the extracted GO biological

process (BP) terms was a *p*-value of less than 0.01 and a FDR of less than 0.05 (Benjamini-Hochberg correction), while the *p*-value threshold for the KEGG pathway enrichment analysis was set to less than 0.05. The results are illustrated in bar charts using the "ggplot" R package. For further functional enrichment analysis of the hub genes, we introduced the "ClueGO"plugin[21] in Cytoscape with the "GO term fusion" function to visualize the pivotal biological processes and pathways involved.

Results

Workflow

As shown in Fig. 1, WGCNA analysis was performed on the combined normalized expression matrix of GSE41177 and GSE79768, and batch effects were removed using the "sva" package. After the co-expression network was constructed, the key gene modules specifically associated with the LAA were detected and the hub genes, functional enrichment and interaction among proteins were subsequently identified. In parallel, we investigated the common DEGs by making comparisons between LAA vs RAA and LAA vs LA-PV. Crucial genes were defined as genes that were identified as both DEGs and modular hub genes, as well as the top-ranked intramodular hub genes identified through the various methods outlined above.

Microarray data processing

The GSE41177 and GSE79768 datasets were downloaded from the GEO database and converted into a standard expression matrix. After preprocessing using the "sva" package to remove batch effects, the top 25% of the most variant probes (13,669 probes) were selected from a total of 64 samples (LAA, $n = 32$; RAA, $n = 13$; LA-PV, $n = 19$) for the WGCNA analysis.

Construction of a weighted co-expression network and identification of modules associated with LAA

The sample clustering dendrogram is shown in Fig. 2. Using a cut-off height of 135, the following three samples were determined to be outliers and were removed from the clusters (shown in a red frame): GSM2102198, GSM2102206 and GSM1005420. The adjacency matrix was obtained based on the approximate scale-free topology criterion (Fig. 3): with a soft-power threshold β of 8 given by the function "sft$powerEstimate," the scale independence could be up to 0.9. Then, the module genes were detected using the TOM matrix. After being merged with a MEDissThres threshold of 0.25, a total of 31 modules were detected (Fig. 4). Interaction-based relationships and eigengene adjacencies are illustrated in Fig. 5A-B, indicating a high level of independence among the modules. After matching modules with clinical traits (Fig. 5C), the highest level of association was found in the pink module (composed of 397 probes) and black module (composed of 419 probes) (MEpink: $r = 0.22$, $p = 9.7\times10^{-6}$; MEblack: $r = 0.27$, $p = 2\times10^{-8}$) using Pearson's correlation analysis (Fig. 5D-E). In addition, the pink module was also found to be negatively correlated with RAA, LA-PV and atrial fibrillation (AF). Therefore, the black and pink modules were considered to be the key modules and were chosen for further analysis.

Selection of candidate hub genes in the key modules

Candidate hub genes in the key modules were filtered out using a gene significance (GS) threshold of > 0.2 and module membership (MM) of > 0.8, and a total of 95 and 104 probes were selected from the pink and black modules, respectively. After re-annotation into official gene symbols using the "hgu133plus2.db" database package, probes with the same symbols were removed. The top 30 hub genes were ranked based on intramodular connectivity (kME) and are shown in Table S1. The expression values of the hub genes are shown in heatmaps in Fig. 6, which indicated a high-degree of tissue specificity and correlation with regard to expression patterns.

Functional enrichment analysis of genes in the key modules

GO and KEGG functional enrichment analyses were conducted on the hub genes in the pink and the black modules. The GO and KEGG terms were extracted for further analysis using the threshold of an adjusted p-value of < 0.05. The GO-BP (biological process) terms and KEGG pathways with the top 15 gene counts in the two modules were filtered out (Fig. 7A-B & 7E-F). Meanwhile, with the "Use GO Term Fusion" function, all similar and highly connected biological processes were merged, and pivotal processes that could be enriched were highlighted (Fig. 7C & 7F). GO-CC (cellular component) and GO-MF (molecular function) analyses were also conducted (data not shown).

The functional analysis of the pink module indicated a relatively loose association with biological processes, which were mainly enriched in serine-type endopeptidase activity, positive regulation of vascular endothelial growth factor production, cornification, Staphylococcus aureus infection, multicellular organismal water homeostasis and endothelial cell differentiation. Nevertheless, cell-cell adhesion, synapse and axon guidance and protein processing were found to be notably involved among the top-ranked processes and pathways of the pink module.

The black module demonstrated a close relationship with several metabolic processes: triglyceride metabolic process, triglyceride biosynthetic process, fatty acid oxidation, carbohydrate biosynthetic process, insulin signaling pathway, regulation of lipid storage peroxisome proliferators-activated receptor (PPAR) signaling pathway, primary alcohol metabolic process and regulation of lipolysis adipocytes. Moreover, processes involved in a response to peptide hormone and amino acid biosynthesis also appeared in the functional enrichment analysis.

Construction of PPI network

The protein-protein interaction (PPI) network of the top 30 hub genes in the two key modules (pink and black) were also constructed. In the pink module, the two highly connected hub genes were filtered out: LRRN4 (leucine-rich repeat neuronal protein 4) and KLK11 (kallikrein related peptidase 11).

In the black module, we identified three genes involved in carbohydrate and triacylglycerol metabolisms that were clustered as top hub genes: GYG2 (glycogenin 2), GPD1 (glycerol-3-phosphate dehydrogenase 1) and DGAT2 (diacylglycerol O-acyltransferase 2). All five genes were selected as crucial genes (I).

DEG analysis, overlap with key modules and selection of crucial genes

Using a cutoff value of |logFC|>1 and FDR of < 0.05, a total of 28 common DEGs were identified between the two sets compared: LAA vs RAA and LAA vs LA-PV (Fig. 9A). Through analyzing gene overlap between the pink and black modules with the DEGs, a relatively larger overlap was found between DEGs and the pink module, compared to that of the black module (17 vs 1; Fig. 9B-C). All common DEGs showed a significantly higher expression level in LAA than in RAA and LA-PV, which was similar to the pink module's pattern of association in the module-trait relationship heatmap (Fig. 5C). All 18 common DEGs were selected as crucial genes (II) for further functional enrichment analysis together with crucial genes (I).

Functional enrichment analysis of crucial genes

We performed another functional enrichment analysis on all crucial genes (Table S2), and 6 of the 23 genes were clustered into nine biological processes (Fig. 10). Four crucial genes, RBP4, GPD1, DGAT2 and GYG2, mainly participated in carbohydrate biosynthetic processes, while the other two genes, ANXA8 and ANXA8L1, were found to be associated with (endo-) peptidase and (phospho-) lipase activities.

Discussion

In this study, we performed a WGCNA analysis to determine the key co-expression modules specifically related to the LAA, which may illuminate mechanisms that underlie metabolomics changes after surgical or interventional LAA closure.

The atrial appendages are considered as remnants from heart development and are divers from the atrial anatomy.[22] The myocytes of the appendages resemble ventricular myocytes rather than atrial myocytes, and form comb-like trabeculae within the appendages.[23] As a critical hormone maintaining fluid-salt balance, atrial natriuretic peptide (ANP), is mainly secreted by the LAA under mechanical stress and excessive volume.[24] In addition, the innervation of autonomic nervous system (ANS) in the ganglionic plexi (GPs) of the peri-LAA epicardial adipose tissue constitute the baroregulatory hub of the left atrial appendage.[25]

As a technique for eliminating the LAA from systemic circulation to prevent thrombus formation, epicardial ligation and endocavitary occlusion affects LAA tissue in different ways. In contrast to total exclusion through surgical intervention, which leads to necrosis and fibrosis of the LAA distal to the point of ligation within 4 weeks,[26,27] endocardial occluders exert a permanent and eccentric stretch force at the LAA landing zone, which ensures optimal compression of the device as recommended.[28] Moreover, a minor peri-occluder leak of less than 5 mm, often ignored in LAAC procedures, may persist after neo-endothelialization and thus cause elevated local wall stress.[29] Therefore, sustained higher stress in LAA tissue from using an occlusion device implies a higher local energy demand and metabolic level.

Changes in local energy demands of LAA are not an isolated phenomenon. Several studies have documented the correlation between LAAC and systemic metabolism, which may be further related to cardiac remodeling and long-term prognosis. It has been postulated that systemic neurohormonal modulations after LAAC can be mediated via two pathways: the natriuretic peptide pathway, which accounts for most short-term effects, and the neurally mediated pathway, which affects afferent fibers within the LAA and injured peri-LAA ganglionated plexi.[2] Meanwhile, alterations of metabolomic profiles after LAAC have also been found, including products of glycolysis, tricarboxylic acid cycle (TCA), urea metabolism,[30] lipidome metabolites[31,32] and amino acids.[33]

In this study, we applied WGCNA analysis as our data mining method to simplify the complicated genetic network under investigation. By combining two microarray datasets after eliminating batch effects, we identified the two gene modules with the highest levels of significant association to the LAA.

The pink module is related to a wide range of biological processes and pathways, involving processes such as endopeptidase activities, synapse formation and axon guidance, cell-cell adhesion and immune response. The hub genes enriched in LRRN4 (leucine-rich repeat neuronal protein), together with other leucine-rich repeat proteins, have been reported to play a crucial role during the process of dilated cardiomyopathy.[34,35] KLK11 (kallikrein 11), which is part of the kallikrein kinin system (KKS) is involved in the pathogenesis of heart failure and has been mainly discussed in relation to genitourinary tumors, while its other functions involving heart tissue remain to be discovered. Intriguingly, annexin A8 (ANXA8), a member of the annexin family that performs multiple functions, including impacts on proliferation, quiescence and colony formation in epithelial cells, while also acting as an indirect inhibitor of the thromboplastin-specific complex,[36] was identified as a hub gene in the LAA.[36] Its effect on cardiac tissue has rarely been studied.

The black module was found to be highly associated with lipid and carbohydrate metabolism. DGAT2, or ladiacylglycerol O-acyltransferase 2, which catalyzes the final step in triglyceride (TG) synthesis, has received increasing recognition for its role in regulating cardiac lipid turnover and affecting cardiac physiology. Partial inhibition of DGAT2 activity increased cardiac fatty acid oxidation, and its co-inhibition with DGAT1 in the heart abrogated TG turnover and protected the heart from high fat diet-induced lipid accumulation with no adverse effects on cardiac function.[37] GPD1 encodes for glycerol-3-phosphate dehydrogenase 1, which is well known to catalyze the reversible conversion of reduced and oxydized nicotine adenine dinucleotide (NADH-NAD+) and is indispensable for energy-yielding processes.[38] An increase in GPD1 translation has been observed in hypertrophic heart induced by pregnancy, which may be mediated by the activation of hypoxia inducible factor-1α (HIF-1α) and PPARγ (the Warburg effect).[39] The adipokine retinol-binding protein-4 (RBP4) in plasma and adipose tissue indicate future cardiometabolic risk related to metabolic syndrome[35,41] and have been reported to be implicated in insulin resistance and vascular inflammation.[42] RBP4 has also been correlated with plasma levels of the mid-regional pro atrial natriuretic peptide (MR-proANP), a biomarker of heart failure[43], as well

as the toll-like receptor 4 and myeloid differentiation primary response gene 88 (TLR4/MyD88) mediated inflammatory pathway, which contributes to both insulin resistance and cardiac hypertrophy.[44]

Overall, the integrated bioinformatics co-expression analysis highlighted the strong relationship between the LAA and innervation, as well as glycolipid metabolism, which may indicate a higher energy demand and neuromodulation effect on the LAA compared with other atrial tissues, such as RAA and LA-PV. However, further experimental verification will still be needed to confirm the mechanism by which these hub genes participate in systemic neurohormonal changes after LAA exclusion.

Study Limitations And Perspectives

This bioinformatic study was based on human heart tissue samples obtained from *in silico* studies stored on the GEO database. The study has provided insights into the probable mechanisms through which the LAA is implicated in neurohormonal changes after ligation or occlusion, since post-procedural samples obtained from the LAA were hardly available due to ethical reasons. More *in vivo* research, including the verification of LAA-occluder-implanted swine models, and the overexpression or knockdown of specific genes, remain to be conducted in the future. In addition, as hypothesized in our discussion and the “LAA HOMEOSTASIS” study,[2] we attributed the association between post-procedural local and systemic metabolomic changes to neural feedbacks. However, since the LAA is well known as an endocrine organ, potential hormonal and molecular modulations need to be further clarified using blood samples.

Conclusion

The bioinformatics approach of WGCNA was used to provide a comprehensive understanding of key genes and pathways in LAA in order to elucidate the specific biological characteristics of the human LAA, and our results provide some novel insights in the underlying on potential mechanisms of neuro-hormonal changes following LAA exclusion procedures.

Declarations

Conflicts of interest:

none.

Funding:

This work was supported by the State Key Program of National Natural Science Foundation of China (grant number 81530015),the Shanghai Science and Technology Committee (grant number 19411963500) Shanghai Jiao Tong University School of Medicine (grant number 2018DGDGJJL02 & JZPI201706), Funding for Clinical Trial of Xinhua Hospital (grant number 19XHCR19D). Hospital Funded

Basic Research, Xinhua Hospital Affiliated to Shanghai Jiao Tong University School of Medicine (grant number 17YJJ09).

References

1. Xu Y-Z, et al. The Modulating Effects of Cardiac Resynchronization Therapy on Myocardial Metabolism in Heart Failure. Pacing Clin Electrophysiol. 2016;39:1404–9.
2. Lakkireddy D, et al. Left Atrial Appendage Closure and Systemic Homeostasis: The LAA HOMEOSTASIS Study. J Am Coll Cardiol. 2018;71:135–44.
3. Susan-Resiga D, et al. Furin is the major processing enzyme of the cardiac-specific growth factor bone morphogenetic protein 10. J Biol Chem. 2011;286:22785–94.
4. Martin RIR, et al. Genetic variants associated with risk of atrial fibrillation regulate expression of PITX2, CAV1, MYOZ1, C9orf3 and FANCC. J Mol Cell Cardiol. 2015;85:207–14.
5. Barabasi A-L, Gulbahce N, Loscalzo J. Network medicine: a network-based approach to human disease. Nat Rev Genet. 2011;12:56–68.
6. Zhang B, Horvath S. A general framework for weighted gene co-expression network analysis. Stat Appl Genet Mol Biol. 2005;4:Article17–7.
7. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
8. Yin L, et al. Identification of novel blood-based HCC-specific diagnostic biomarkers for human hepatocellular carcinoma. Artif Cells Nanomedicine Biotechnol. 2019;47:1908–16.
9. Yang W, et al. Co-expression Network Analysis Identified Key Proteins in Association With Hepatic Metastatic Colorectal Cancer. Proteomics - Clin Appl. 2019;13:1–11.
10. Wang HZ, et al. Coexpression network analysis identified that plakophilin 1 is associated with the metastasis in human melanoma. Biomed Pharmacother. 2019;111:1234–42.
11. Zhang H, et al. Co-Expression Network Analysis Identified Gene Signatures in Osteosarcoma as a Predictive Tool for Lung Metastasis and Survival. J Cancer. 2019;10:3706–16.
12. Niemira M, et al. Molecular signature of subtypes of non-small-cell lung cancer by large-scale transcriptional profiling: Identification of key modules and genes by weighted gene co-expression network analysis (WGCNA). *Cancers (Basel)*. 12, (2020).
13. Ding M, Li F, Wang B, Chi G, Liu H. A comprehensive analysis of WGCNA and serum metabolomics manifests the lung cancer-associated disordered glucose metabolism. J Cell Biochem. 2019;120:10855–63.
14. Wang JD, et al. Prediction of competing endogenous RNA coexpression network as prognostic markers in AML. Aging (Albany NY). 2019;11:3333–47.
15. Chen R, et al. Identification of biomarkers correlated with hypertrophic cardiomyopathy with co-expression analysis. J Cell Physiol. 2019;234:21999–2008.

16. Yeh Y-H, et al. Region-specific gene expression profiles in the left atria of patients with valvular atrial fibrillation. Hear Rhythm. 2013;10:383–91.
17. Tsai FC, et al. Differential left-to-right atria gene expression ratio in human sinus rhythm and atrial fibrillation: Implications for arrhythmogenesis and thrombogenesis. Int J Cardiol. 2016;222:104–12.
18. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey J. D. The SVA package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28:882–3.
19. Chen S, et al. Identification of crucial genes in abdominal aortic aneurysm by WGCNA. *PeerJ* 2019, 1–20 (2019).
20. Shannon P, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.
21. Mlecnik B, Galon J, Bindea G. Automated exploration of gene ontology term and pathway networks with ClueGO-REST. Bioinformatics. 2019;35:3864–6.
22. DeSimone CV, et al. A Review Of The Relevant Embryology, Pathohistology, And Anatomy Of The Left Atrial Appendage For The Invasive Cardiac Electrophysiologist. J Atr Fibrillation. 2015;8:1129.
23. Al-Saady NM, Obel OA, Camm AJ. Left atrial appendage: structure, function, and role in thromboembolism. Heart. 1999;82:547–54.
24. Kappagoda CT, Linden RJ, Snow HM. The effect of distending the atrial appendages on urine flow in the dog. J Physiol. 1972;227:233–42.
25. Arora R. Recent insights into the role of the autonomic nervous system in the creation of substrate for atrial fibrillation: implications for therapies targeting the atrial autonomic nervous system. Circ Arrhythm Electrophysiol. 2012;5:850–9.
26. Ellis CR, Byrd JM, Scalf SL. Ischemic necrosis of the left atrial appendage at autopsy 4 weeks following epicardial suture ligation via a sub-xiphoid approach (LARIAT). J Interv Card Electrophysiol. 2015;43:99–100.
27. Holmes DR, Reddy VY. Left Atrial Appendage and Closure: Who, When, and How. Circ Cardiovasc Interv. 2016;9:e002942.
28. Meincke F, Schmidt-Salzmann M, Kreidel F, Kuck K-H, Bergmann MW. New technical and anticoagulation aspects for left atrial appendage closure using the WATCHMAN® device in patients not taking warfarin. EuroIntervention. 2013;9:463–8.
29. Behnes M, et al. –LAA Occluder View for post-implantation Evaluation (LOVE)–standardized imaging proposal evaluating implanted left atrial appendage occlusion devices by cardiac computed tomography. BMC Med Imaging. 2016;16:25.
30. Sattler K, et al. Occlusion of left atrial appendage affects metabolomic profile: focus on glycolysis, tricarboxylic acid and urea metabolism. Metabolomics. 2017;13:1–9.
31. Yücel G, et al. Percutaneous Closure of Left Atrial Appendage significantly affects Lipidome Metabolism. Sci Rep. 2018;8:1–13.

32. Fastner C, et al. Interventional left atrial appendage closure affects the metabolism of acylcarnitines. Int J Mol Sci. 2018;19:1–12.
33. Rusnak J, et al. Interventional left atrial appendage closure may affect metabolism of essential amino acids and bioenergetic efficacy. Int J Cardiol. 2018;268:125–31.
34. Li R, et al. Leucine-rich repeat neuronal protein 4 (LRRN4) potentially functions in dilated cardiomyopathy. Int J Clin Exp Pathol. 2017;10:9925–33.
35. Brody MJ, Lee Y. The Role of Leucine-Rich Repeat Containing Protein 10 (LRRC10) in Dilated Cardiomyopathy. Front Physiol. 2016;7:337.
36. Jiang X, et al. Annexin A8 (ANXA8) regulates proliferation of porcine endometrial cells via Akt signalling pathway. Reprod Domest Anim. 2019;54:3–10.
37. Roe ND, Handzlik MK, Li T, Tian R. The Role of Diacylglycerol Acyltransferase (DGAT) 1 and 2 in Cardiac Metabolism and Function. Sci Rep. 2018;8:4983.
38. Oeljeklaus S, Schummer A, Mastalski T, Platta HW, Warscheid B. Regulation of peroxisome dynamics by phosphorylation. Biochim Biophys Acta. 2016;1863:1027–37.
39. Soñanez-Organis JG, et al. HIF-1α and PPARγ during physiological cardiac hypertrophy induced by pregnancy: Transcriptional activities and effects on target genes. Gene. 2016;591:376–81.
40. Zachariah JP, et al. Prospective Relation of Circulating Adipokines to Incident Metabolic Syndrome: The Framingham Heart Study. J Am Heart Assoc. 2017;6:e004974.
41. Smith U, Kahn BB. Adipose tissue regulates insulin sensitivity: role of adipogenesis, de novo lipogenesis and novel lipids. J Intern Med. 2016;280:465–75.
42. Yang Q, et al. Serum retinol binding protein 4 contributes to insulin resistance in obesity and type 2 diabetes. Nature. 2005;436:356–62.
43. Yagmur E, et al. Elevated MR-proANP plasma concentrations are associated with sepsis and predict mortality in critically ill patients. J Transl Med. 2019;17:415.
44. Gao W, et al. Retinol-Binding Protein 4 Induces Cardiomyocyte Hypertrophy by Activating TLR4/MyD88 Pathway. Endocrinology. 2016;157:2282–93.

Figures

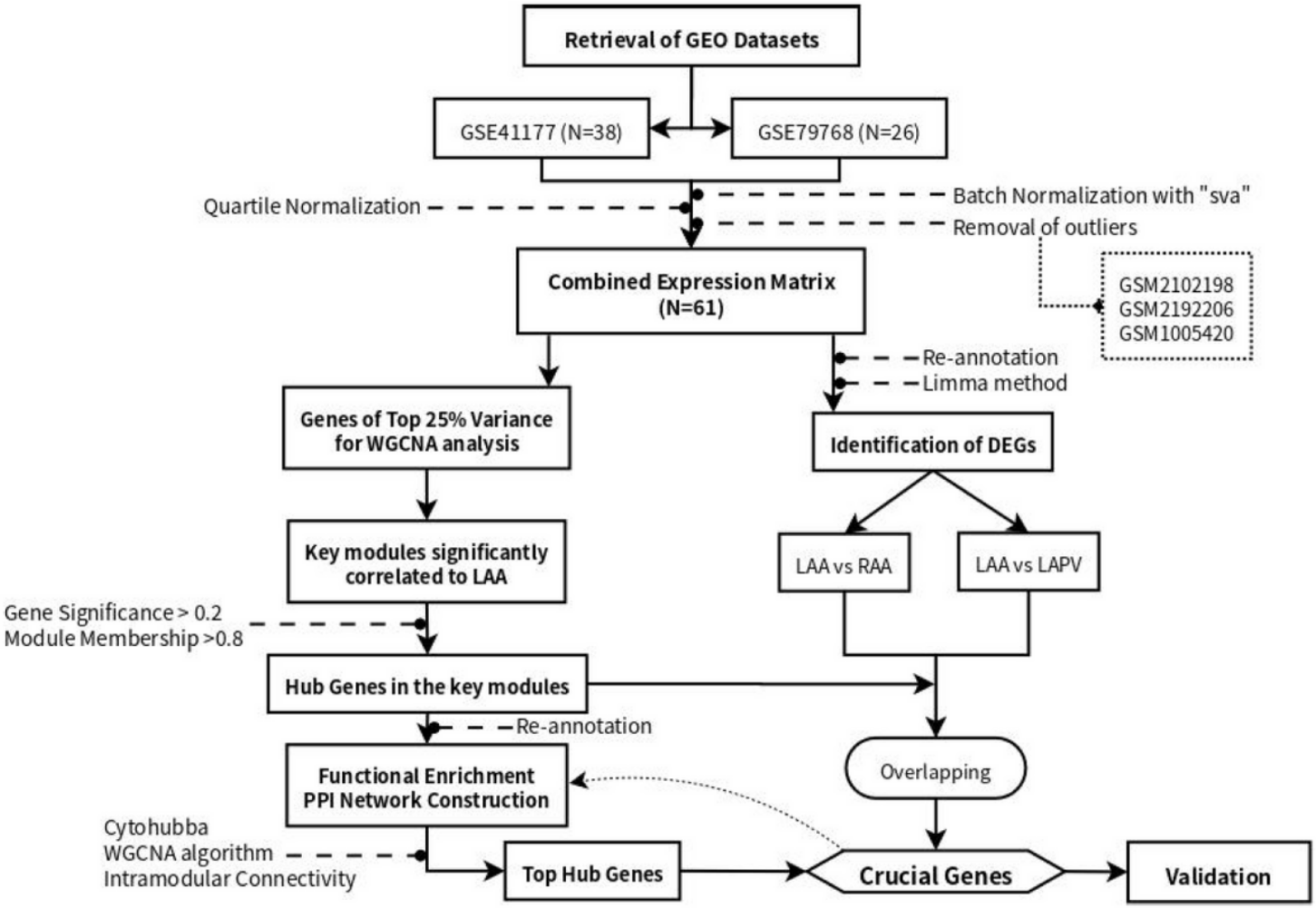

Figure 1

Workflow of the entire study. GEO, Gene Expression Omnibus; DEGs, differentially expressed genes; LAA, left atrial appendage; RAA, right atrial appendage; LAPV, left atrial junction; WGCNA, weighted gene co-expression network analysis; PPI, protein-protein interaction.

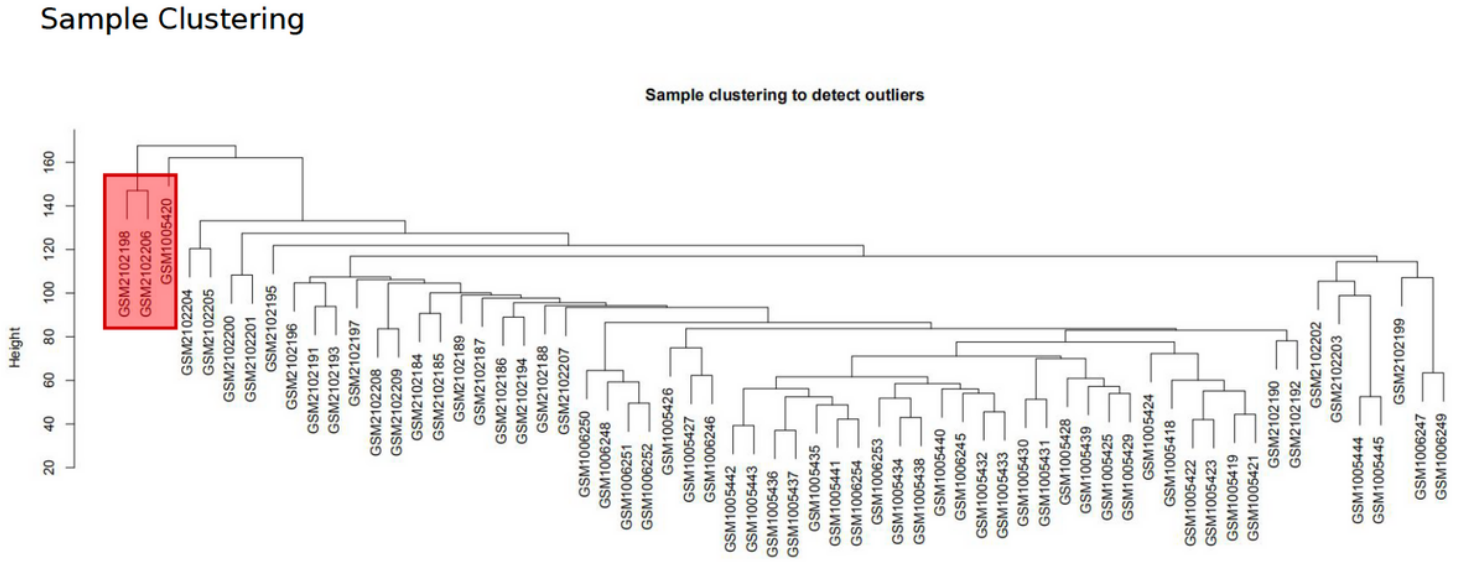

Figure 2

Cluster analysis conducted using the flashClust package based on batch-normalized data to detect outliers. Data from three samples shown in the red frame were eliminated before the WGCNA analysis was conducted using a cut-off height of 135.

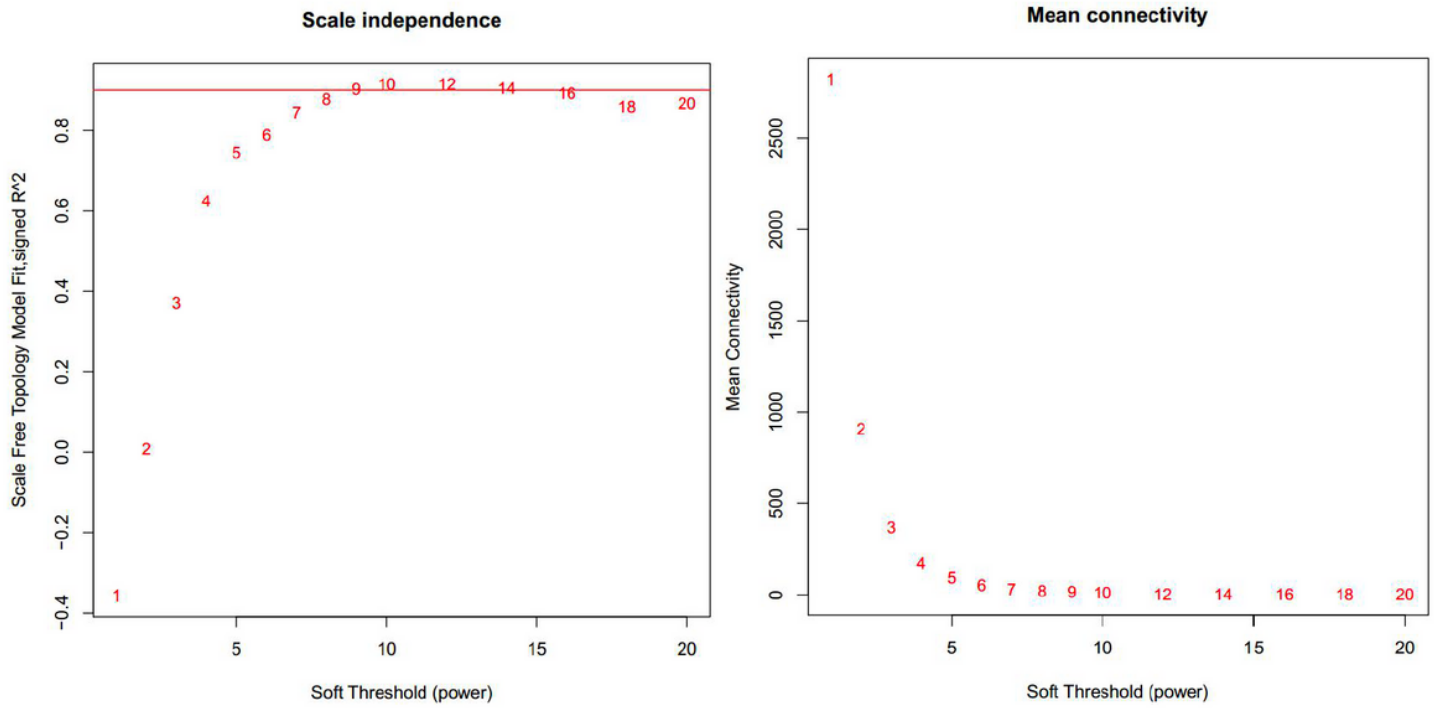

Figure 3

Network topology for determining the soft-thresholding power. The numbers shown on the plots indicate the corresponding soft thresholding power (β). The approximate scale-free topology can be attained at a soft-thresholding power of 8.

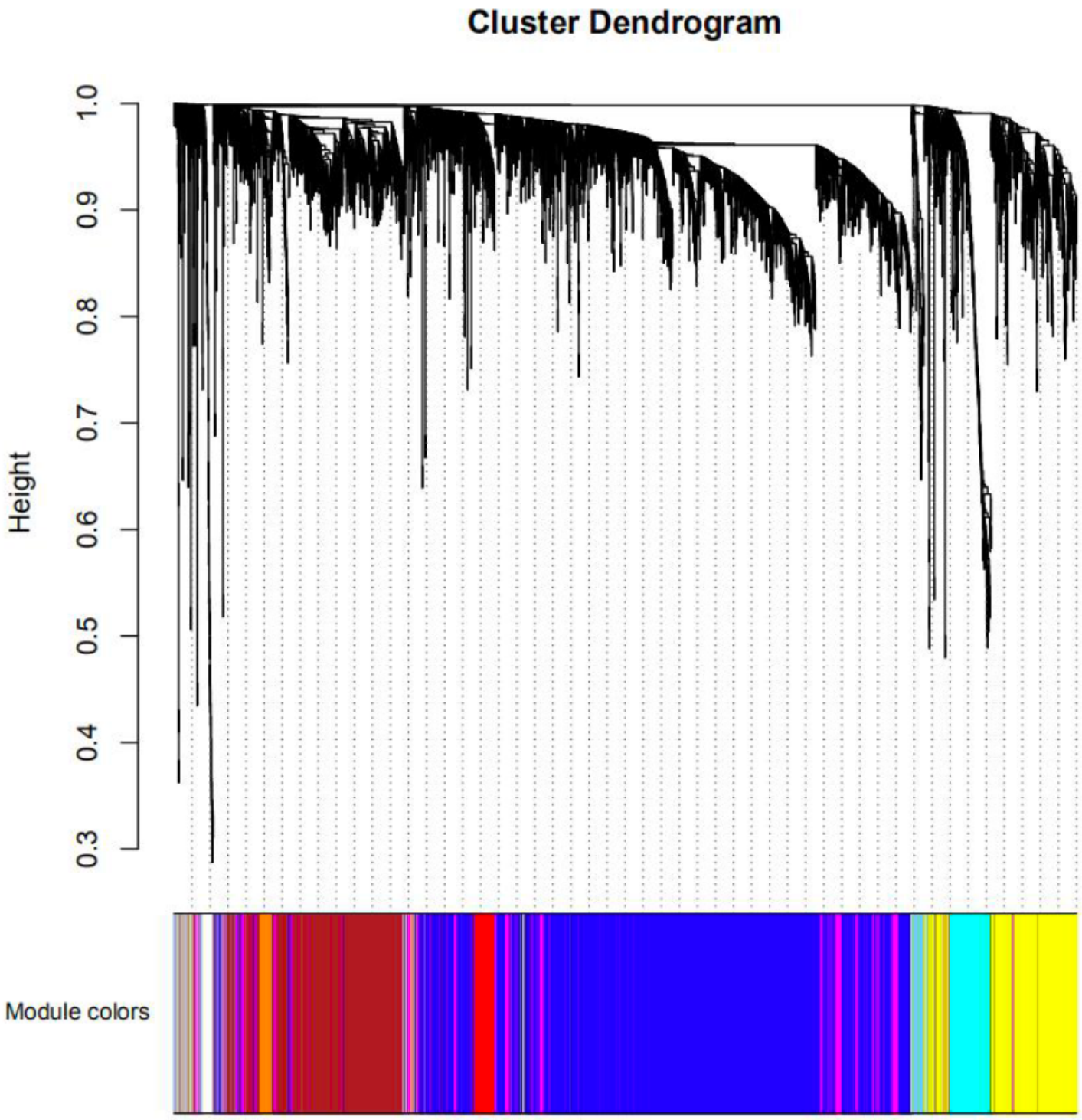

Figure 4

Construction of co-expression modules using the WGCNA package in R. Hierarchical cluster analysis was conducted to detect co-expression clusters with corresponding color assignments. Each branch represents a single gene, and each color below represents one co-expression module.

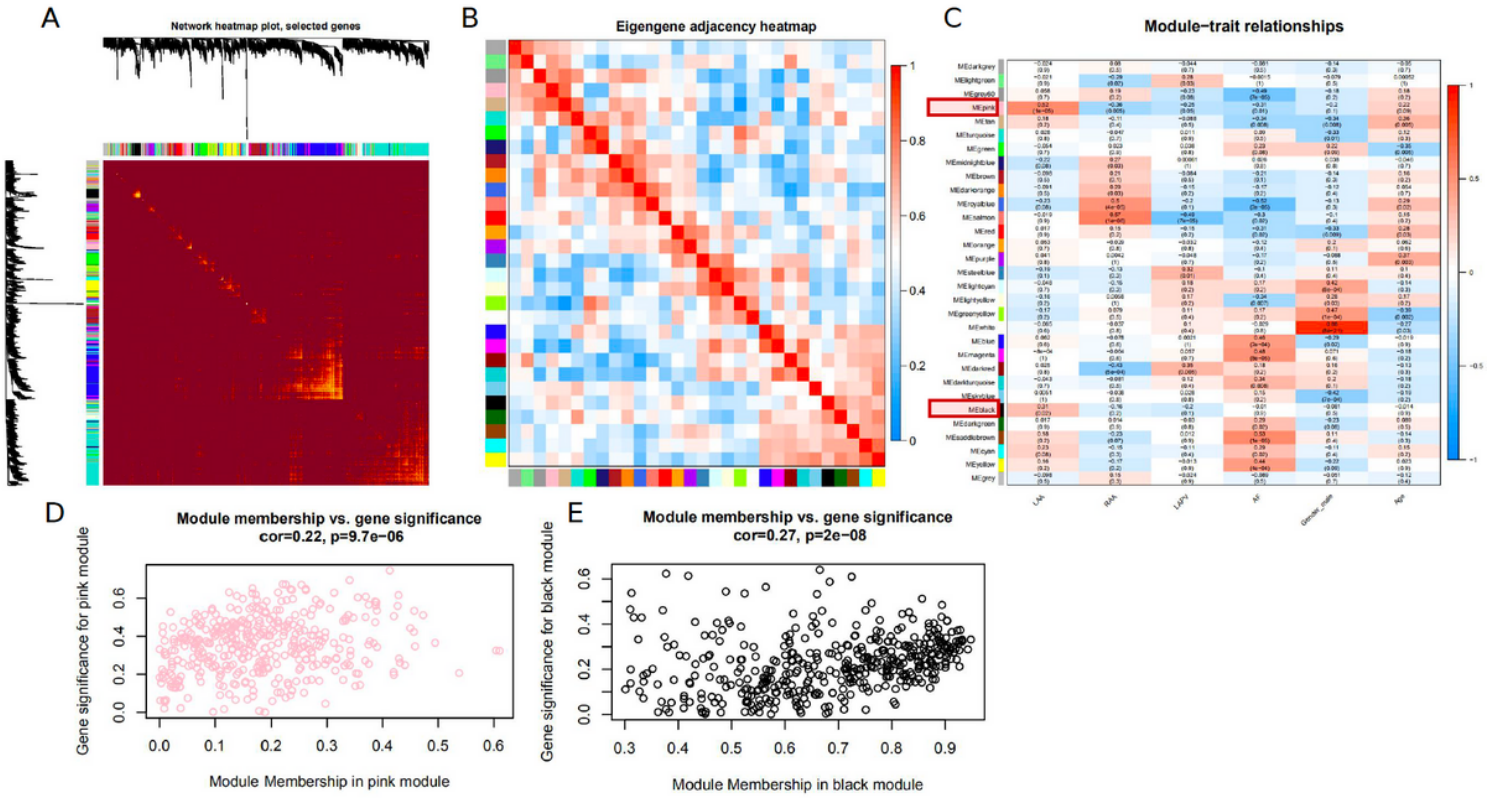

Figure 5

(A) Association analysis of the modules. Different colors in the horizontal and vertical axes represent different modules. The brightness of the yellow color in the middle of the figures represent the connectivity of different modules. There was no significant dissimilarity of association (1-TOM) found among the different modules, showing a high-degree of independence among these modules. TOM, topological overlap matrix. (B) Heat map plot displaying the adjacencies of eigengenes in the different modules. (C) Heatmap displaying the correlations and significant differences between gene modules and clinical traits. Correlations are displayed in the rectangle, while significant differences are displayed in brackets. MEpink and MEblack were found to be significantly correlated to LAA. (D-E) A scatterplot of Gene Significance (GS) for LAA vs. Module Membership (MM) in the pink and black modules, respectively. LAA, left atrial appendage.

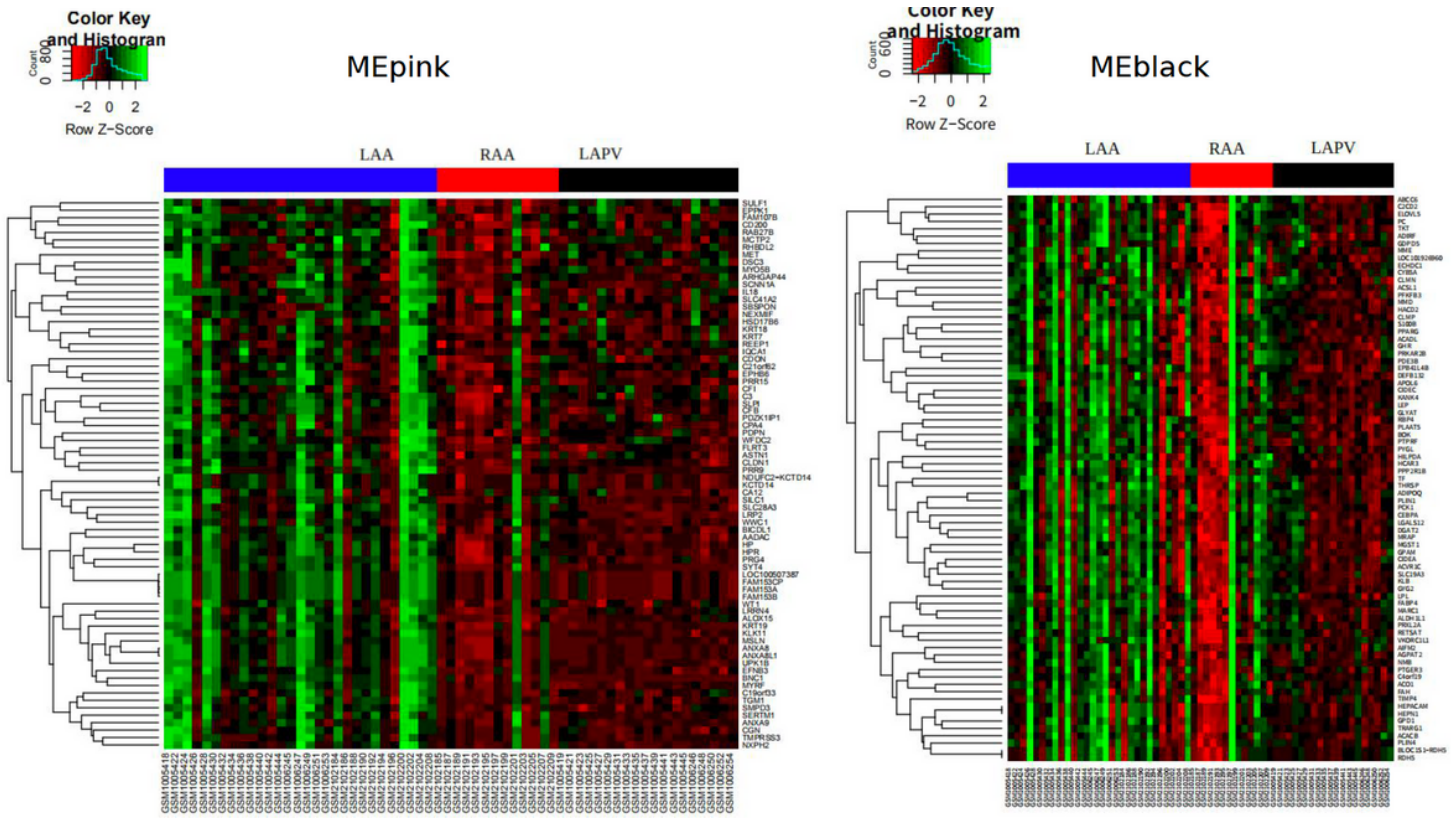

Figure 6

Expression values of hub genes in the key modules shown in the heat maps.

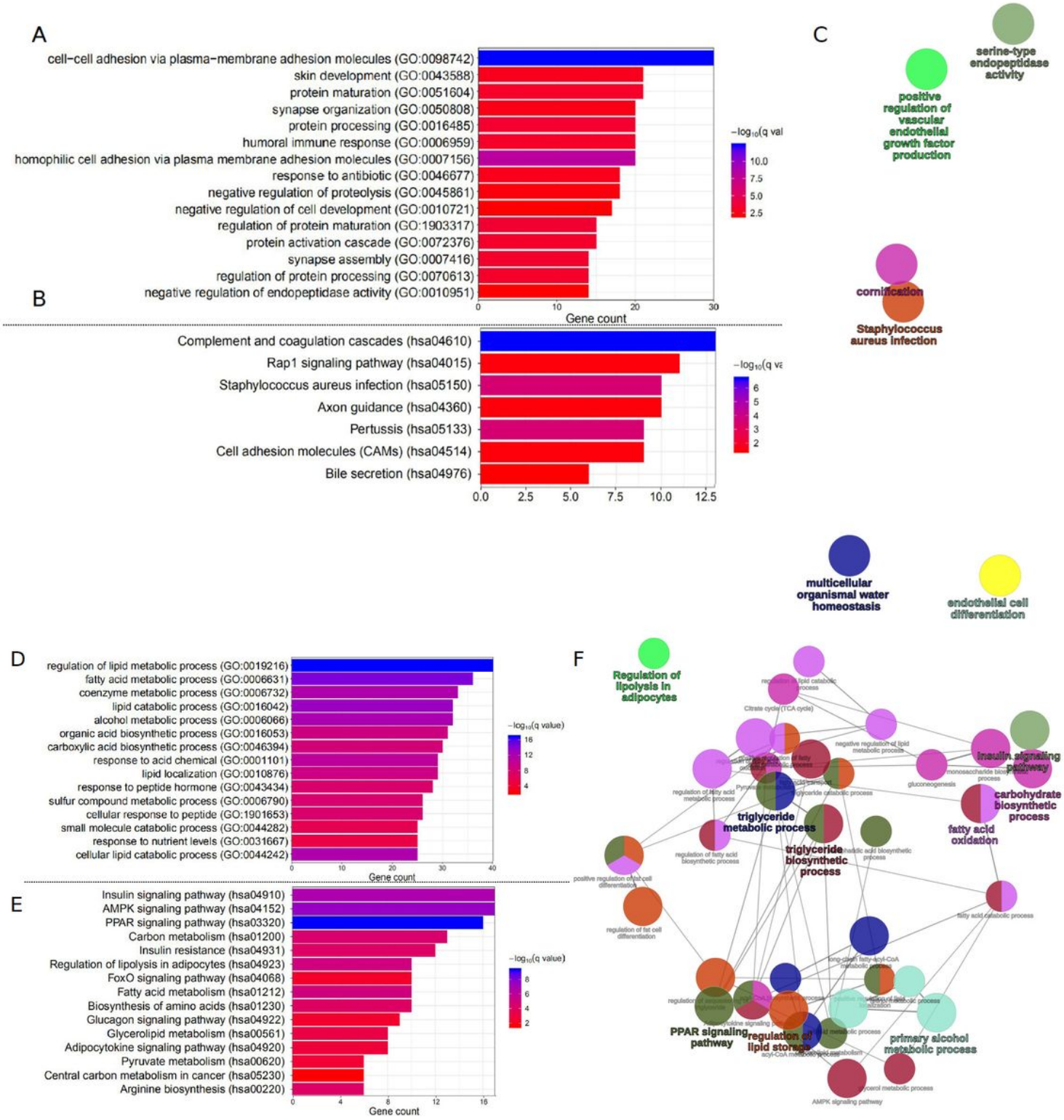

Figure 7

Functional enrichment analysis of the pink and black module genes. (A) GO-BP analysis in pink module. (B) KEGG analysis in the pink module. (C) Functional enrichment analysis of the pink module using the“ClueGO” plugin and “GO term fusion” function. The color of each bar corresponds with the p value. The top 15 terms were selected based on gene counts. The color of each bar corresponds to the adjusted p value using Benjamini-Hochberg correction. Functional enrichment analysis of the pink and black

module genes. (D) GO-BP analysis in pink module. (E) KEGG analysis in pink module. (F) Functional enrichment of black module with “ClueGO” plugin and “GO term fusion” function. The color of each bar corresponds with p value. Top 15 terms for gene counts were selected. The color of each bar corresponds with adjusted p value with Benjamini-Hochberg correction. GO, gene ontology; BP, biology process; KEGG, Kyoto Encyclopaedia of Genes and Genome.

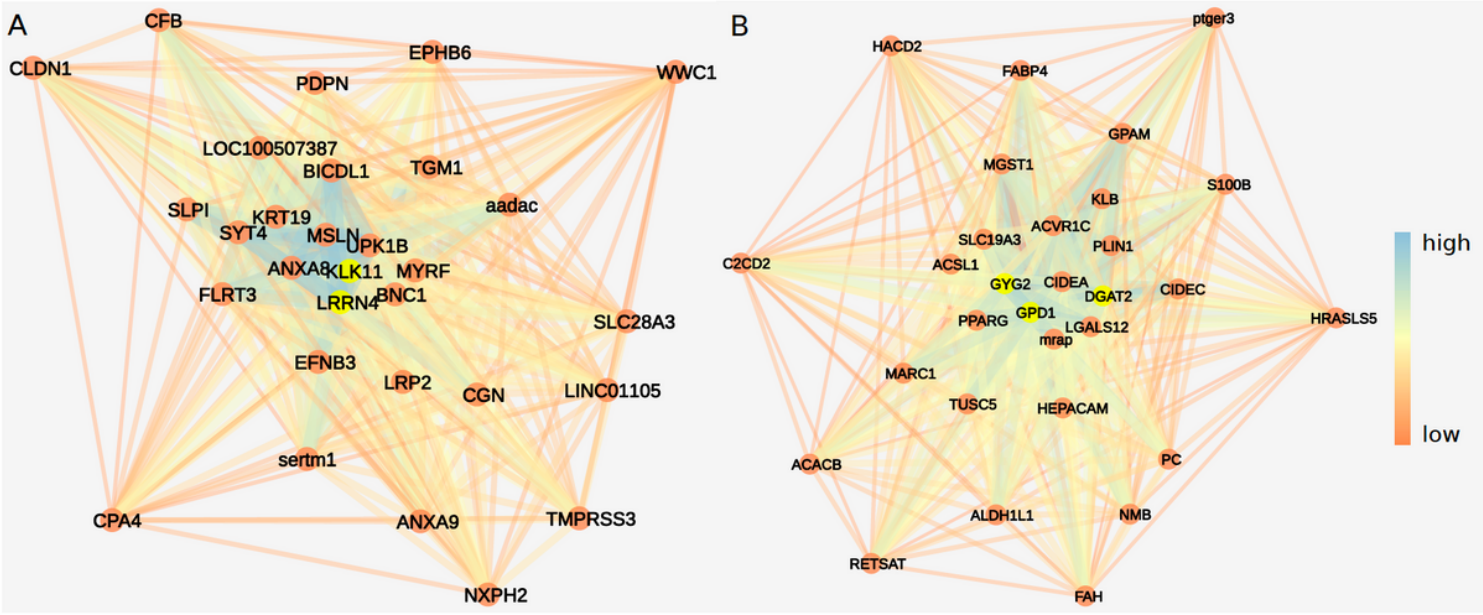

Figure 8

The networks of the top 10 hub genes based on connectivity in the pink (A) and black B) modules. Map edge size and color correspond to the weighted betweenness of the hub genes. Node degrees are shown using different colors. Nodes with a red border indicate genes with the highest rank in each module. KLK11 (kallikrein related peptidase 11) and DGAT2 (diacylglycerol O-acyltransferase 2) were filtered out using cytoHubba, while LRRN4 (leucine rich repeat neuronal 4), GYG2 (glycogenin 2) and GPD1 (glycerol-3-phosphate dehydrogenase 1) were selected based on intramodular connectivity and the WGCNA algorithm. WGCNA, weighted gene co-expression network analysis.

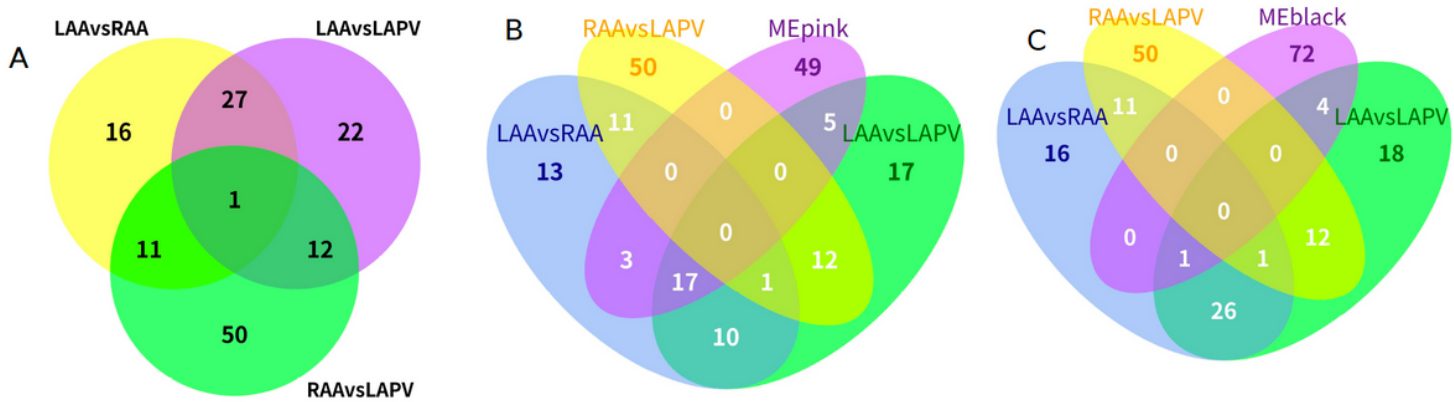

Figure 9

Overlapping between the DEGs and key modules. (A) Venn diagram showing the three pairs of comparisons in the different regions; (B) overlap with the pink module; (C) overlap with the black module. DEGs, differentially expressed genes; LAA, left atrial appendage; RAA, right atrial appendage; LAPV, left atrium-pulmonary vein junction; MEpink, pink module; MEblack, black module.

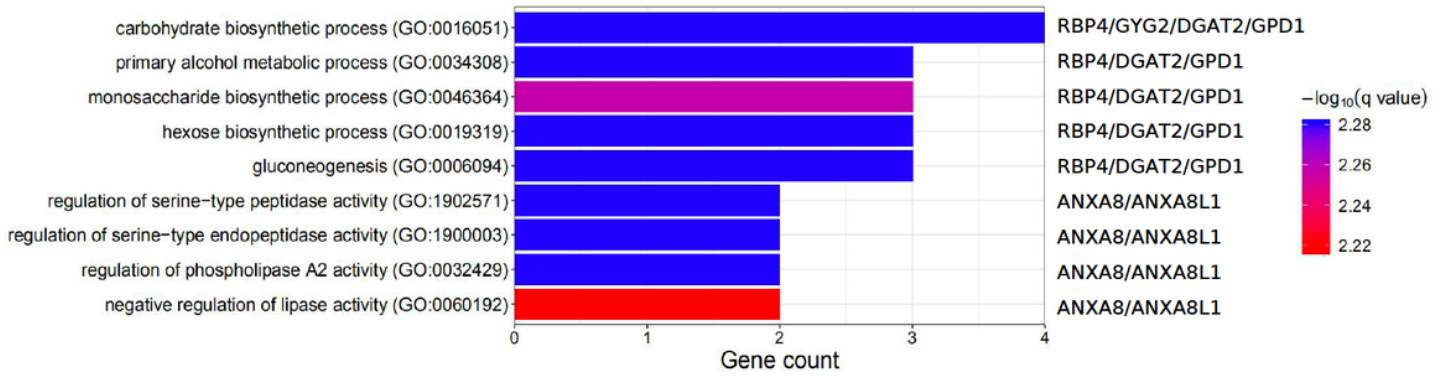

Figure 10

Functional enrichment analysis of crucial genes. The color of each bar corresponds to the adjusted p value using Benjamini-Hochberg correction.

Supplementary Files

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

- SupplementalFigureandTables.pdf