Identification of potential key pathways, genes and circulating markers in the development of intracranial aneurysm based on weighted gene co-expression network analysis

Guoia Du, Dangmurenjiafu Geng, Kai Zhou, Yandong Fan, Riqing Su, Qingjiu Zhou, Bo Liu and Serick Duysenbi

Department of Neurosurgery, The First Affiliated Hospital of Xinjiang Medical University, Urumqi, China

ABSTRACT

Background: Intracranial aneurysm (IA) is a disease resulted from weak brain control, characterized by local expansion or dilation of brain artery. This study aimed to construct a gene co-expression network by Weighted Gene Correlation Network Analysis (WGCNA) to explore the potential key pathways and genes for the development of IA.

Method: Six IA-related gene expression data sets were downloaded from the Gene Expression Omnibus (GEO) database for identifying differentially expressed genes (DEGs). WGCNA was used to identify modules associated with IA. Functional enrichment analysis was used to explore the potential biological functions. ROC analysis was used to find markers for predicting IA.

Results: Purple, greenyellow and yellow modules were significantly associated with unruptured intracranial aneurysms, while blue and turquoise modules were significantly associated with ruptured intracranial aneurysms. Functional modules significantly related to IA were enriched in Ribosome, Glutathione metabolism, CAMP signalling pathway, Lysosome, Glycosaminoglycan degradation and other pathways. CD163, FCEREG, FPR1, ITGAM, NLRC4, PDG, and TYROBP were up-regulated in ruptured intracranial aneurysms and serum, these genes were potential circulating markers for predicting IA rupture.

Conclusions: Potential IA-related key pathways, genes and circulating markers were identified for predicting IA rupture by WGCNA analysis.

Introduction

Intracranial aneurysm (IA) is a disease resulted from weak brain control, characterized by local expansion or dilation of brain artery; typically, cystic aneurysm is the most commonly seen, which is the leading cause inducing spontaneous subarachnoid haemorrhage (SAH) [1]. The overall prevalence rate of unruptured intracranial aneurysm (UIA) among populations with no common disease is 3.2%, with the average age of 50 years old [2,3]. In recent years, IA has been increasingly detected due to the frequent application of imaging technique in clinic [4,5]. Although some IAs remain stable and unbroken, part of them may encounter rupture and endanger life [6]. Previous studies showed the mortality of SAH induced by ruptured intracranial aneurysm (RIA) is approximately 30–40%, moreover, half of survivors may suffer disability [7,8]. Consequently, there is still an urgent need to investigate the pathogenesis of IA for prevention and treatment for this disease, and a reliable circulating marker for predicting RIA can help physicians to prevent the severe consequences induced by RIA.

The Weighted Gene Co-expression Network Analysis (WGCNA) has been verified as an effective method to identify the co-expression modules and hub genes [9–16]. Through WGCNA, Genes with the similar expression pattern into a module can be clustered and the correlation of the module with the phenotype can be analyzed [17,18]. In the present study, WGCNA was employed to identify UIA-related and RIA-related modules for exploring potential IA-related key pathways and hub genes. In addition, some hub genes were also differentially expressed in blood and may be potential circulating markers for predicting RIA.

Materials and methods

Data collection and pre-processing

In this study, all data were derived from the Gene Expression Omnibus (GEO) database (PMID: 23193258), and processed using the Rv3.6.0 (https://www.r-project.org/). Six IA-related data sets were downloaded from the database, including GSE6551, GSE13353, GSE26969, GSE75436, GSE106520, and GSE36791. Among them, the GSE6551, GSE13353, GSE26969, GSE75436, GSE106520, and GSE36791 datasets were raw data from the GPL570 platform. These datasets covered 33 IA wall tissue samples and 27 normal arterial wall controls; The gene expression profile...
Differentially expressed gene analysis and clustering analysis

After the above-mentioned pre-processing of the several data sets, the limma package was used to screening DEGs between IAs and controls. The DEGs between UIAs and controls, the DEGs between RIAs and controls were screened in the combined data set, respectively. \( p \) Value \(<0.05\) considered to be significant. Bidirectional hierarchical clustering analysis using the expression values of DEGs based on the Euclidean distance was performed, and the results were presented as the heat maps.

**The weighted gene co-expression network analysis**

WGCNA is a comprehensive set of R functions used to perform analysis of various aspects of weighted correlation networks [18]. WGCNA was performed using the common DEGs between UIA compared to control and RIA compared to control. The aim of WGCNA was to identify the correlation with gene modules and the three phenotypes (control, UIA, and RIA). First, weighted value of person coefficient was used to calculate the correlation any two genes. Then, the hierarchical clustering tree was constructed through the Pearson coefficients between genes, in which the different branches indicated different gene modules, while different colours stood for different modules. The result of correlation with gene modules and the three phenotypes were showed as a
heat map. A module with high correlation with the phenotype is the candidate module for further investigation.

**Enrichment analysis for gene modules**

Gene Ontology (GO) is a database constructed by the Association for Gene Ontology. GO annotations can be classified as three categories, including molecular function (MF), biological process (BP), and cellular components (CC) [22]. Kyoto Encyclopaedia of Genes and Genomes (KEGG) is a comprehensive database integrating the genomic, chemical and systemic functional information, among which, the KEGG Pathway is specifically used to store gene pathway information among different species [23,24]. To explore the biological functions of the candidate gene modules, enrichment analyses including GO and KEGG pathways enrichment analysis were performed using clusterProfiler package [25]. p Value < .05 and q value < 0.05 were set as threshold.

**Identification of the Hub genes and circulating markers**

Through WGCNA, gene significance (GS) is defined as the correlation a gene with phenotype. The module membership (MM): \( MM(i) = \text{cor}(x, i, \text{ME}) \) is defined to measure the importance of a gene in the module. In the present study, a gene with GS > 0.2 and MM > 0.8 was defined as hub gene among the candidate gene modules. In order to find a reliable biomarker for predicting IA, the expression of hub gene was investigated whether the differential expression in blood tissue was consistent with the differential expression in arterial wall tissue. Furthermore, hub genes with consistent differential expression were carried out receiver operating characteristic curve (ROC) analysis to identify the potential circulating markers to diagnose UIA and predict RIA. The ROC analysis was performed using pROC package [26]. Through WGCNA analysis, key pathways, genes related to IA and potential circulating markers were identified that may predict aneurysm rupture.

**Results**

**DEG and clustering analysis**

According to the p value < .05, we found that 783 genes were significantly up-regulated and 1097 genes were significantly down-regulated in UIA (Figure 2(A)). While 711 genes were significantly up-regulated and 1020 genes were significantly down-regulated in RIA (Figure 2(B)). The most up-regulated 50 genes and the most down-regulated 50
genes were used to perform cluster analysis, the results showed that the expression patterns of these 100 DEGs can basically distinguish IA and normal control (Figure 2(C,D)).

The results of WGCNA

In the present study, 2559 common DEGs in RIA and UIA were used to perform WGCNA (Figure 3(A)). Prior to WGCNA, we had carried out sample clustering analysis, in which two outlier samples were discovered and eliminated (Figure 3(B)).

With regard to the WGCNA parameter setting, $R^2 = 0.85$, $\beta = 7$, the minimum gene number in gene module was 30, and the cutting height was set at 0.75. Finally, 12 modules (Figure 3(C)) were identified and each module was assigned with a unique colour (Figure 3(D)). Three modules (purple, greenyellow, and yellow) were markedly correlated with the UIA, while two modules (blue and turquoise) were evidently related to RIA (Figure 3(E)). The genes within modules were significantly correlated with the modules (Figure 4).

Figure 3. Identification of gene modules with WGCNA. (A) Venn diagram for common differentially expressed genes. (B) Sample clustering tree. (C) Determination of adjacency function parameter $\beta$. (D) Identification of gene modules. (E) Correlation heat map of gene modules and phenotypes.
Enrichment analysis

In order to explore the biological functions of these five candidate gene modules, GO and KEGG pathway enrichment analysis were performed. Enrichment results (Figure 5) in GO biological process (BP) showed that, the formation of IA might be related to neutrophil degranulation, neutrophil activation involved in immune response, neutrophil mediated immunity, and neutrophil activation. In addition, the results of KEGG pathway enrichment analysis revealed that, modules related to UIA were involved in the pathways such as

![Figure 4. Scatter plot of gene and module correlation within the module of interest. (A) Blue module. (B) Yellow module. (C) Purple module. (D) Turquoise module. (E) Greenyellow module.](image-url)
Glutathione metabolism, Biosynthesis of amino acids, cAMP signalling pathway, Ras signalling pathway, cGMP-PKG signalling pathway, Glycosaminoglycan degradation and Carbon metabolism. While modules related to RIA were involved in the pathways such as Ribosome, Spliceosome, Lysosome, Phagosome, Toll-like receptor signalling pathway, Osteoclast...
differentiation and NOD-like receptor signalling pathway. These results indicate that the mechanism of IA formation and rupture may be different.

**Identification of the hub genes and IA biomarkers**

Among the above five modules, 101 hub genes were identified according to GS $>0.2$ and MM $>0.8$. The blood gene expression profile of GSE106520 indicated that some hub genes were also differentially expressed in blood in IAs. Figure 6(A) is the heat map for the expression of the 101 hub genes in these three data sets. The expression of 24 hub genes in blood were consistent with the differential expression of artery wall tissues in RIA (Figure 6(B)). While, in the blood gene expression profile GSE36791 from UIA, these hub genes were not significant differential expression, except ALOX5 (Figure 6(C)). The ROC curves showed that CD163, FCEREG, FPR1, ITGAM, NLRC4, PDG and TYROBP1 were potential circulating markers for predicting RIA with high AUC values (Figure 6(D)). ROC curve for ALOX5 indicated that the diagnostic value for UIA may be not well (Figure 6(E)).

**Discussion**

Multiple surveys have indicated the morbidity of IA is estimated to be as high as 10%. Generally, medium and small IAs are asymptomatic prior to rupture. However, patients
with IA will encounter severe SAH in the case of IA rupture [27,28], while SAH is a devastating subtype of stroke with high risk of neurological deficit and even death [29,30]. Therefore, to identify the key potential pathways, genes and circulating markers in the development of IA is very meaningful. In the present study, the DEGs related to IA were identified for carrying out WGCNA. The gene modules correlated with IA formation and rupture were also identified using WGCNA. According to the results, a total of 12 gene modules were identified. Interestingly, we found that three modules (purple, green-yellow and yellow) were markedly correlated with UIA, while the blue and turquoise modules were evidently correlated with RIA. Such results suggested that IA formation and progress might be resulted from different functional modules.

In order to prevent the progression of IA, the mechanism of IA rupture should be further studied. In our study, five interested gene modules were elected to carry out the functional enrichment analysis according to the correlation between phenotype and disease. And the result of enrichment analysis shows that modules related to UIA and RIA were involved in different biological process and pathway, which demonstrated that the mechanism of IA formation and rupture may be different. Previous studies have shown, IA is related to immune response, regulation of immune response, lysome, cGMP PKG signalling pathway, Toll-like receptor signalling and lysosome pathway [31–33], which were also confirmed in our study. Recent studies have demonstrated, lysosomes not only play a central role in cell decomposition, but also participate in metabolism, membrane repair and cell death [34]. And in our study, lysosome pathway was enriched in module related to RIA, which suggests the degradation of the components of the unruptured aneurysm wall may be a process leading to rupture. In addition, many studies showed that IA is also closely related to immune/inflammatory response [31,32,35,36]. Inflammation is one of the crucial players in the pathophysiology of IA, inflammation and immune response could potentially contribute to the formation of IA [37]. However, the specific mechanism of immune/inflammation in the formation and rupture of IA is not clear. Additionally, we also found Ribosome, Glutathione metabolism, cAMP signalling pathway and Glycosaminoglycan degradation were probably related to the formation and rupture of IA, but more detailed mechanism of IA rupture needs to be further explored.

Moreover, we identified 101 hub genes from five interested gene modules, which subsequently were used to ROC curve analysis. Then the results of ROC curves showed that CD163, FCEREG, FPRL, ITGAM, NLRC4, PDG and TYROBP were potential circulating markers for predicting RIA. Among them, CD163 is a scavenger receptor for hemoglobin-haptoglobin complex phagocytosis [38], it is highly expressed in degenerated and ruptured IA walls [39]. The upregulation of FPRL can induce the synthesis of neutrophils mediated by Mac-1 through fpr related Gi protein and β-arrestin signalling pathway, thus aggravating the development of abdominal aortic anerysm (AAA) [40], which suggest it may also participate in the development of IA. In addition, NLRC4 has been found to play an important role in human autoinflammation, TYROBP was also studied in inflammatory diseases [41], but whether these two genes were related to inflammation of IA rupture needs further study. The identification of these genes can help IA patients to predict whether there is a danger of rupture in aneurysms, and remind doctors to intervene as soon as possible, which is of great significance.

Though our study provided new insights in IA, the potential key pathways and genes were identified based on bioinformatics method and need to be validated in molecular experiment. Therefore, it is not clear whether these genes are causal or merely markers for UIA or RIA.

Conclusion

There has no reliable circulating markers to precisely diagnose IA at present. In the present study, the hub genes related to IA formation and progress were identified, their predictive value for IA were explored. Gene CD163, FCEREG, FPRL1, ITGAM, NLRC4, PDG, and TYROBP are up-regulated in the both arterial wall tissues and blood tissues, and these gene may potentially serve as the circulating markers for predicting RIA.

Acknowledgement

The authors would like to thank the Doc Chenxi Liu, the Biomedic Technology Co., Ltd and Life-Ontology Biological Technology Co., Ltd for assisting with bioinformatics analysis.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This study was funded by grants of National Natural Science Foundation Programme of China (NSFC Programme) [81760222 and 31360209].

Data availability statement

The datasets generated during and analyzed during the current study are available in the NCBI-GEO Dataset repository, [https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE36791, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE106520, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE75436, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE26969, https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE36791].

Author contributions

Guojia Du designed the study and wrote the paper. Dangmurenjiafu Geng and Kai Zhou and Yandong Fan managed the literature searches and analyses. Riqing Su, Qingjiu Zhou, Bo Liu and Serick Duyensbi wrote the first draft of the manuscript. All authors contributed to and approved the final manuscript.
References

[1] Wardlaw JM, White PM. The detection and management of unruptured intracranial aneurysms. Brain. 2000;123(2):205–221.

[2] Vlak MH, Algra A, Brandenburg R, et al. Prevalence of unruptured intracranial aneurysms, with emphasis on sex, age, comorbidity, country, and time period: a systematic review and meta-analysis. Lancet Neurol. 2011;10(7):626–636.

[3] Chan DY, Abrego JM, Cheung TC, et al. Screening for intracranial aneurysms? Prevalence of unruptured intracranial aneurysms in Hong Kong Chinese. JNS. 2016;24(5):1245–1249.

[4] Etminan N, Rinkel GJ. Unruptured intracranial aneurysms: development, rupture and preventive management. Nat Rev Neurol. 2016;12(12):699–713.

[5] Brown RD, Jr., Broderick JP. Unruptured intracranial aneurysms: epidemiology, natural history, management options, and familial screening. Lancet Neurol. 2014;13(4):393–404.

[6] Frosen J, Tulamo R, Paetau A, et al. Saccular intracranial aneurysm: pathology and mechanisms. Acta Neuropathol. 2012;123(6):773–786.

[7] Nieuwkamp DJ, Setz LE, Algra A, et al. Changes in case fatality of aneurysmal subarachnoid haemorrhage over time, according to age, sex, and region: a meta-analysis. Lancet Neurol. 2009;8(7):635–642.

[8] Stegmayr B, Eriksson M, Asplund K. Declining mortality from subarachnoid hemorrhage: changes in incidence and case fatality from 1985 through 2000. Stroke. 2004;35(9):2057–2063.

[9] Bo L, Wei B, Wang Z, et al. Screening of critical genes and microRNAs in blood samples of patients with ruptured intracranial aneurysms by bioinformatic analysis of gene expression data. Med Sci Monit. 2017;23:4518–4525.

[10] Guo X, Xiao H, Guo S, et al. Identification of breast cancer mechanism based on weighted gene coexpression network analysis. Cancer Gene Ther. 2017;24(8):333–341.

[11] Huang H, Zhang Q, Ye C, et al. Identification of prognostic markers of high grade prostate cancer through an integrated bioinformatics approach. J Cancer Res Clin Oncol. 2017;143(2):2571–2579.

[12] Liu X, Hu AX, Zhao JL, et al. Identification of key gene modules in human osteosarcoma by co-expression analysis weighted gene co-expression network analysis (WGCNA). J Cell Biochem. 2017;118(11):3953–3959.

[13] Ma C, Lv Q, Teng S, et al. Identifying key genes in rheumatoid arthritis by weighted gene co-expression network analysis. Int J Rheum Dis. 2017;20(8):971–979.

[14] Tang Y, Ke ZP, Peng YG, et al. Co-expression analysis reveals key gene modules and pathway of human coronary heart disease. J Cell Biochem. 2018;119(2):2102–2109.

[15] Giulietti M, Occhipinti G, Principato G, et al. Identification of candidate miRNA biomarkers for pancreatic ductal adenocarcinoma by weighted gene co-expression network analysis. Cell Oncol (Dordr). 2017;40(2):181–192.

[16] Giulietti M, Occhipinti G, Principato G, et al. Weighted gene co-expression network analysis reveals key genes involved in pancreatic ductal adenocarcinoma development. Cell Oncol (Dordr). 2016;39(4):379–388.

[17] Presson AP, Sobel EM, Papp JC, et al. Integrated weighted gene co-expression network analysis with an application to chronic fatigue syndrome. BMC Syst Biol. 2008;2:95.

[18] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.

[19] Jeannet P. [Reflections on the notion of the "central masturbation fantasy" of adolescence]. Psychiatr Enfant. 1992;35(1):25–32.

[20] Leek JT, Johnson WE, Parker HS, et al. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882–883.

[21] Tutino VM, Poppenberg KE, Jiang K, et al. Circulating neutrophil transcriptome may reveal intracranial aneurysm signature. PLoS One. 2018;13(1):e0191407.

[22] Gene Ontology C Gene Ontology Consortium: going forward. Nucleic Acids Res. 2017;45(Database issue):D1049–56.

[23] Kanehisa M, Sato Y, Kawashima M, et al. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016;44(D1):D457–D462.

[24] Zhang JD, Wiemann S. KEGGgraph: a graph approach to KEGG PATHWAY in R and bioconductor. Bioinformatics. 2009;25(11):1470–1471.

[25] Yu G, Wang LG, Han Y, et al. ClusterProfiler: an R package for comparing biological themes among gene clusters. OMICS. 2012;16(5):284–287.

[26] Robin X, Turck N, Hainard A, et al. pROC: an open-source package for R and S-+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12:77.

[27] Rinkel GJ. Natural history, epidemiology and screening of unruptured intracranial aneurysms. Rev Neurol (Paris). 2008;164(10):781–786.

[28] Brown RD. Unruptured intracranial aneurysms. Semin Neurol. 2010;30(5):537–544.

[29] Etminan N. Aneurysmal subarachnoid hemorrhage-status quo and perspective. Transl Stroke Res. 2015;6(3):167–170.

[30] Cheng C, Jiang L, Yang Y, et al. Effect of APOE gene polymorphism on early cerebral perfusion after aneurysmal subarachnoid hemorrhage. Transl Stroke Res. 2015;6(6):446–450.

[31] Wang W, Li H, Yu L, et al. Aberrant expression of IncRNAs and mRNAs in patients with intracranial aneurysm. Oncotarget. 2017;8(2):2477–2484.

[32] Kleineuoff R, Verweij BH, van der Vlies P, et al. RNA sequencing analysis of intracranial aneurysm walls reveals involvement of lysosomes and immunoglobulins in rupture. Stroke. 2016;47(5):1286–1293.

[33] Kukl MI, Hakkinen SK, Frosen J, et al. Upregulated signaling pathways in ruptured human saccular intracranial aneurysm wall: an emerging regulatory role of Toll-like receptor signaling and nuclear factor-κB, hypoxia-inducible factor-1A, and ETS transcription factors. Neurosurgery. 2011;68(6):1667–1675.

[34] Serrano-Puebla A, Boya P. Lysosomal membrane permeabilization in cell death: new evidence and implications for health and disease. Ann. N. Y. Acad. Sci. 2016;1371(1):30–44.

[35] Pera J, Korostynski M, Krzyszczowski T, et al. Gene expression profiles in human ruptured and unruptured intracranial aneurysms: what is the role of inflammation? Stroke. 2010;41(2):224–231.

[36] Ollikainen E, Tulamo R, Lehti S, et al. Smooth muscle cell foam cell formation, apolipoproteins, and ABCA1 in intracranial aneurysms: implications for lipid accumulation as a promoter of aneurysm wall rupture. J Neuropathol Exp Neurol. 2016;75(7):689–699.

[37] Xu Z, Li H, Song J, et al. Meta-analysis of microarray-based expression profiles to identify differentially expressed genes in intracranial aneurysms. World Neurosurg. 2017;99:668–668 e7.

[38] Kristiansen M, Graversen JH, Jacobsen C, Sonne O, et al. Identification of the haemoglobin scavenger receptor. Nature. 2001;409(6817):198–201.

[39] Ollikainen E, Tulamo R, Kaitainen S, Sonne O, et al. Macrophage infiltration in the saccular intracranial aneurysm wall as a response to locally lysed erythrocytes that promote degeneration. J Neuropathol Exp Neurol. 2018;77(10):890–903.

[40] He L, Fu Y, Deng J, et al. Deficiency of FAM3D (family with sequence similarity 3, member D), A novel chemokine, attenuates neutrophil recruitment and ameliorates abdominal aortic aneurysm development. Arterioscler Thromb Vasc Biol. 2018;38(7):1616–1631.

[41] Pelham CJ, Agrawal DK. Emerging roles for triggering receptor expressed on myeloid cells receptor family signaling in inflammatory diseases. Expert Rev Clin Immunol. 2014;10(2):243–256.