European Review for Medical and Pharmacological Sciences 2023; 27: 4876-4882

Identification of hub genes correlated with diabetic retinopathy *via* bioinformatics methods

J.-H. XIONG[1], J.-L. CHEN[2], J.-Y. LIANG[1], F.-F. ZHANG[1], S.-M. CHENG[1]

[1]School of Chinese Medicine, Jiangxi University of Chinese Medicine, Nanchang, China
[2]Department of TCM Ophthalmology, Jiangshan's Hospital of Traditional Chinese Medicine, Jiangshan, China

Jianghao Xiong and Jiali Chen are co-first authors

Abstract. – OBJECTIVE: The aim of this study was to identify the hub genes and uncover the molecular mechanisms of diabetic retinopathy (DR).

MATERIALS AND METHODS: We used the Gene Expression Omnibus (GEO) dataset GSE60436 in our study. After screening for differentially expressed genes (DEGs), we performed gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analysis. Subsequently, a protein-protein interaction (PPI) network was constructed using the Search Tool for Retrieval of Interacting Genes (STRING) database and visualized using the Cytoscape software. Finally, we identified 10 hub genes by cytoHubba plugin.

RESULTS: A total of 592 DEGs were identified, including 203 up-regulated genes and 389 down-regulated genes. The DEGs were mainly enriched in visual perception, photoreceptor outer segment membrane, retinal binding, and PI3K-Akt signaling pathway. By constructing a protein-protein interaction (PPI) network, 10 central genes were finally identified, including *CNGA1, PDE6G, RHO, ABCA4, PDE6A, PDE6B, NRL, RPE65, GUCA1B* and *AIPL1*.

CONCLUSIONS: *CNGA1, PDE6G, RHO, ABCA4, PDE6A, PDE6B, NRL, RPE65, GUCA1B*, and *AIPL1* may be potential biomarkers and therapeutic targets for DR.

Key Words:
Diabetic retinopathy, Bioinformatics, Molecular mechanisms, Biomarkers.

Introduction

Diabetes is a common endocrine disease mainly caused by glucose metabolism disorders. The prevalence of diabetes and the number of patients with diabetes are increasing rapidly. The International Diabetes Federation estimates[1] that 1 in 10 adults will be diagnosed with diabetes by 2040, and by 2045, there will be 629 million people with diabetes worldwide. The mortality of diabetes will be higher than that of HIV/AIDS, tuberculosis, and malaria combined with about 5 million people dying each year. Diabetic retinopathy (DR) is one of the most common microvascular complications of diabetes. In China, its prevalence is 24.7-37.5%, and the prevalence of DR is 50-70% in patients with a disease course of 10-15 years. About 2% of patients with diabetes onset in 15 years are completely blind as a result of DR. Its occurrence and development greatly influence people's life quality.

The course of diabetes, blood glucose level, blood pressure, blood lipid, and other clinical factors have a certain influence on the occurrence and development of DR. However, there are great individual differences in the pathogenesis and severity during the onset of diabetes. Some patients with diabetes persist for a long time and have lax blood sugar control but still do not develop DR, while some patients develop DR Rapidly. This difference may be caused by gene polymorphism. In recent years, an increased understanding of the safety and biodistribution of gene delivery vectors has been achieved, and it has been demonstrated that stem cells and gene therapy can reverse degenerative retinopathy. Drugs involve regeneration of atrophic or damaged retinal tissue, neurotrophic factors, immune regulation, and new therapies to replace mutated genes. Along with the progress of science and technology, genetic intervention also gradually increased. Current studies[2-4] on gene therapy focus on related gene drugs against pigment epithelium derived factor, blood vessels inhibition, endothelial inhibition, vascular endothelial growth factor and inhibiting endothelial cell proliferation and formation of new blood vessels, which achieve the aim of clinical prevention and treatment of DR. In this study, we aim to identify the hub genes and uncover the molecular mechanisms of DR.

Corresponding Author: Shaomin Cheng, MD; e-mail: gedrg@stu.cuz.edu.cn

Materials and Methods

Screening for Differentially Expressed Genes (DEGs)

GEO (available at: www.ncbi.nlm.nih.gov/geo) database contains a large amount of gene expression data, mainly gene microarray data and high-throughput sequencing data, which is easily accessible and cost-free to use[5]. GEO2R (available at: www.ncbi.nlm.nih.gov/geo/geo2r/) is a tool to perform differential analysis of gene expression data based on the GEO database, from which we can obtain differentially expressed genes (DEGs)[6]. In this study, GSE60436 dataset, which contains 6 diabetic retinopathy samples and 3 control samples, was used to acquire DEGs based GEO2R tool[7]. Adjusted p-value < 0.05 and absolute values of $\log_2$FC for genes > 2.0 were considered as DEGs.

Gene Functional Enrichment Analysis

Gene Ontology (GO) is a tool to annotate genes and analyze their biological processes and functions, and Kyoto Encyclopedia of Genes and Genomes (KEGG) is a database of signaling pathways. GO functional enrichment analysis of DEGs in three modules of molecular function (MF), cellular composition (CC) and biological process (BP), and KEGG pathway enrichment analysis were performed using the Clusterprofiler package version 4.0 (The R Foundation for Statistical Computing, Vienna, Austria) to analyze the role of DEGs in metabolic processes in the organism[8]. Adjusted p-value < 0.05 was considered statistically significant.

Protein-Protein Interaction (PPI) Network

The Search Tool for the Retrieval of Interacting Genes (STRING; available at: http://string-db.org) (version 11.5) is a database that can be applied to predict protein-protein interactions (PPI), which was used to construct the PPI interaction network[9]. Then, we employed Cytoscape software version 3.9.0 (available at: https://cytoscape.org) to visualize the PPI network[10]. Moreover, plug-in cytoHubba was used to identify the hub genes of the PPI network[11].

Statistical Analysis

R software version 4.0.2 (available at: https://www.r-project.org/) was used to process the experimental data. Measurement data were expressed as X±S. Independent samples and paired samples were tested by independent sample t-test. $p < 0.05$ was considered statistically significant.

Results

Identification of DEGs

A total of 592 DEGs were identified, including 203 upregulated genes and 389 downregulated genes according to the GSE60436 dataset. As presented in the volcano plot (Figure 1), red dots represent upregulated genes and green dots represent downregulated genes.

GO and KEGG Enrichment Analysis

According to the results of GO enrichment analysis (Figure 2), in biological process (BP) category, the DEGs were enriched in visual perception, phototransduction, detection of visible light, detection of light stimulus, detection of abiotic stimuli, detection of external stimuli, sensory system develop-

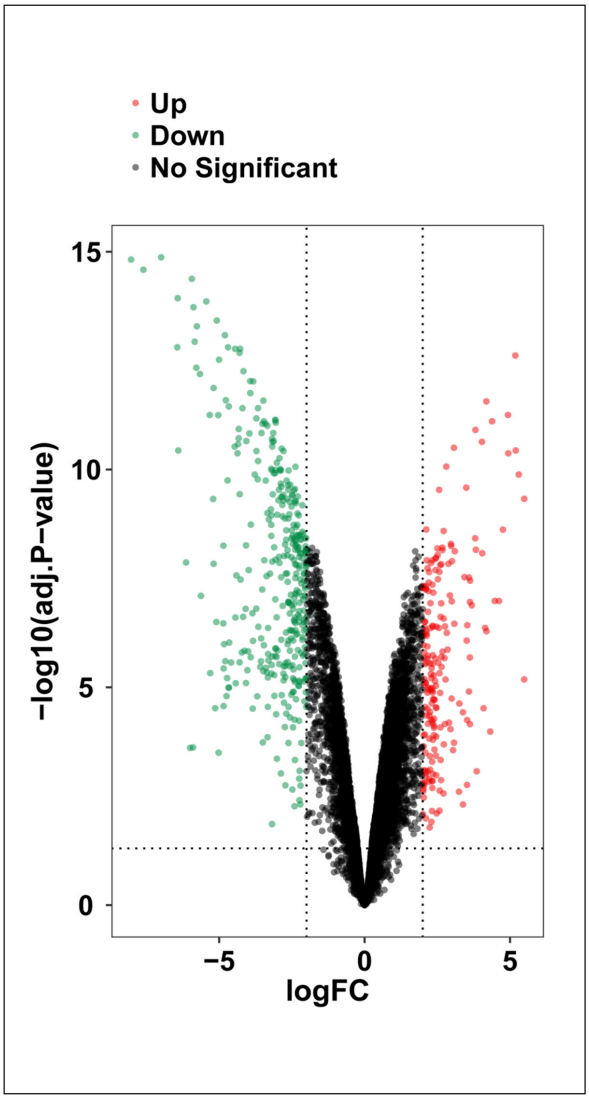

Figure 1. Identification of DEGs by the volcano plot.

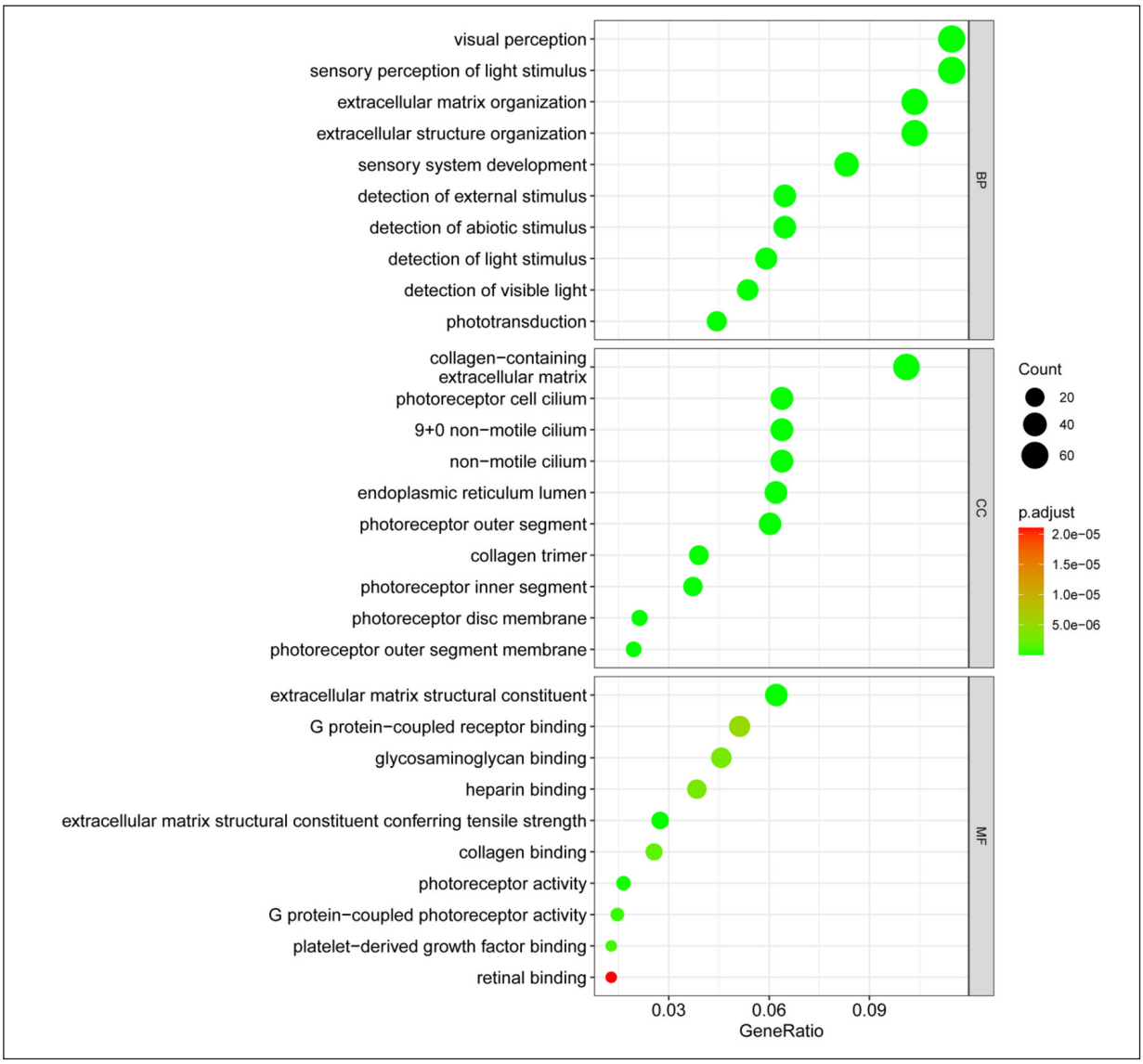

Figure 2. Goes analysis of DEGs.

ment, extracellular structure organization, extracellular matrix organization and sensory perception of light stimulus. As for cellular composition (CC), the DEGs were significantly enriched in photoreceptor outer segment membrane, photoreceptor disc membrane, photoreceptor inner segment, collagen trimer, photoreceptor outer segment, endoplasmic reticulum lumen, non-motile cilium, 9+0 non-motile cilium, photoreceptor cell cilium and collagen-containing extracellular matrix. For CC, the DEGs were enriched in retinal binding, platelet-derived growth factor binding, G protein-coupled photoreceptor activity, photoreceptor activity, collagen binding, extracellular matrix structural constituents conferring tensile strength, heparin binding, glycosaminoglycan binding and G protein-coupled receptor binding. As for KEGG pathway analysis (Figure 3), the most significant pathways of DEGs were phototransduction, extracellular matrix (ECM)-receptor interaction, protein digestion and absorption, PI3K-Akt signaling pathway, arrhythmogenic right ventricular cardiomyopathy, hypertrophic cardiomyopathy, focal adhesion, dilated cardiomyopathy, malaria, phagosome, synaptic vesicle cycle, human papillomavirus infection, cardiac muscle contraction, GABAergic synapse, small cell lung cancer, arginine and proline metabolism, retinol metabolism, AGE-RAGE signaling pathway in diabetic complications, bladder cancer and p53 signaling pathway.

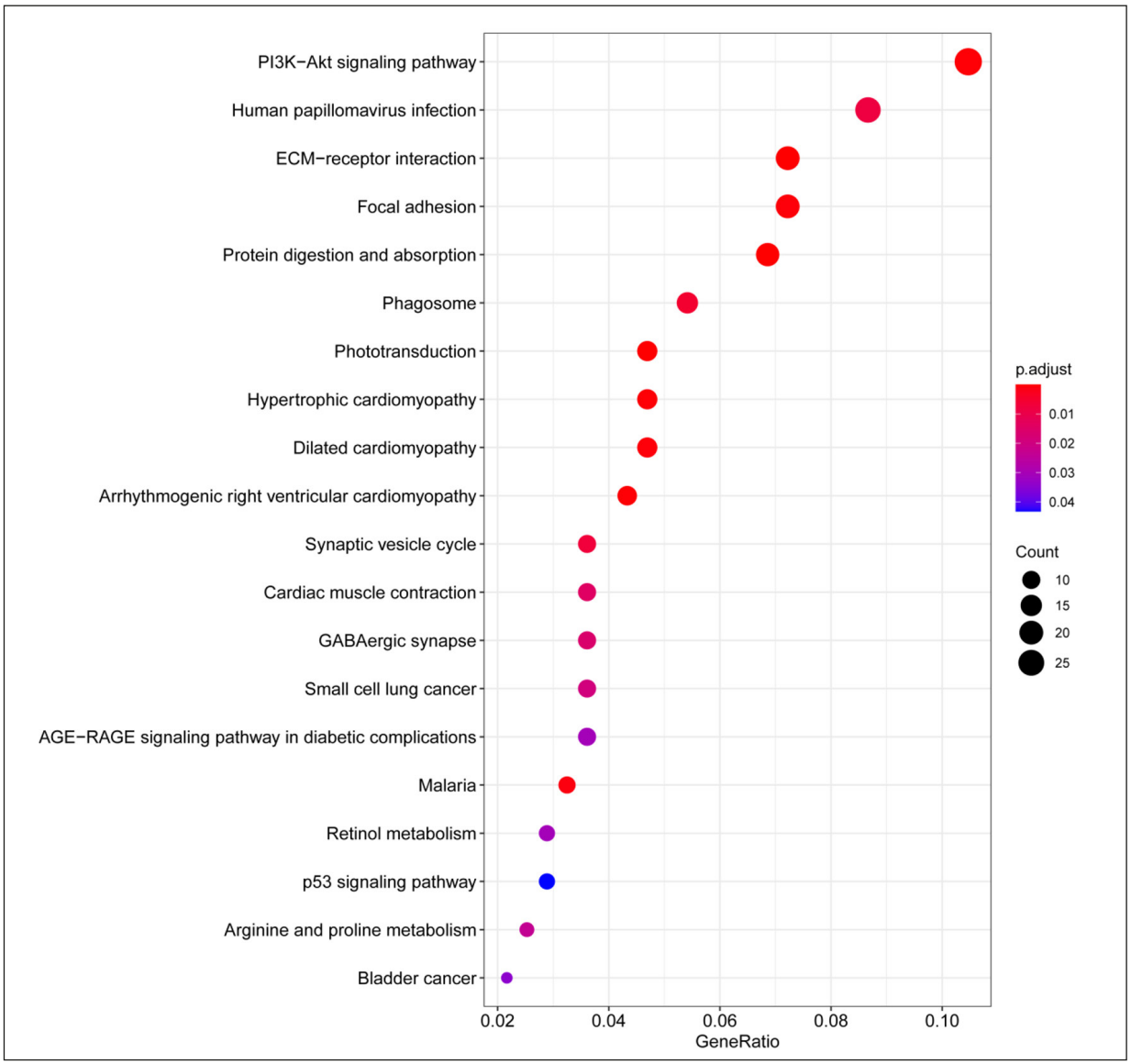

Figure 3. KEGG analysis of DEGs.

Construction of PPI Network and Selection of Hub Genes

There were a total of 3,016 edges and 520 nodes in the PPI network (Figure 4). The greener the color of the gene in the network, the less connected the gene is to other genes. Conversely, the redder the color, the more connected the gene is to other genes. Subsequently, by applying the Matthews correlation coefficient (MCC) algorithm of plug-in cytoHubba, we screened the top 10 hub genes, including *CNGA1*, *PDE6G*, *RHO*, *ABCA4*, *PDE6A*, *PDE6B*, *NRL*, *RPE65*, *GUCA1B* and *AIPL1* (Figure 5).

Discussion

According to statistical forecasts, the total global healthcare expenditure on diabetes is expected to increase to US $802 billion by 2040. According to the Guidelines on Type 2 Diabetes in China[12], the prevalence of diabetes in Chinese adults is 10.9%, and the incidence tends to be younger, with the prevalence of diabetes under 40 years old as high as 5.9%, and the proportion of undiagnosed diabetes is high. Diabetic retinopathy is a progressive and lifelong disease.

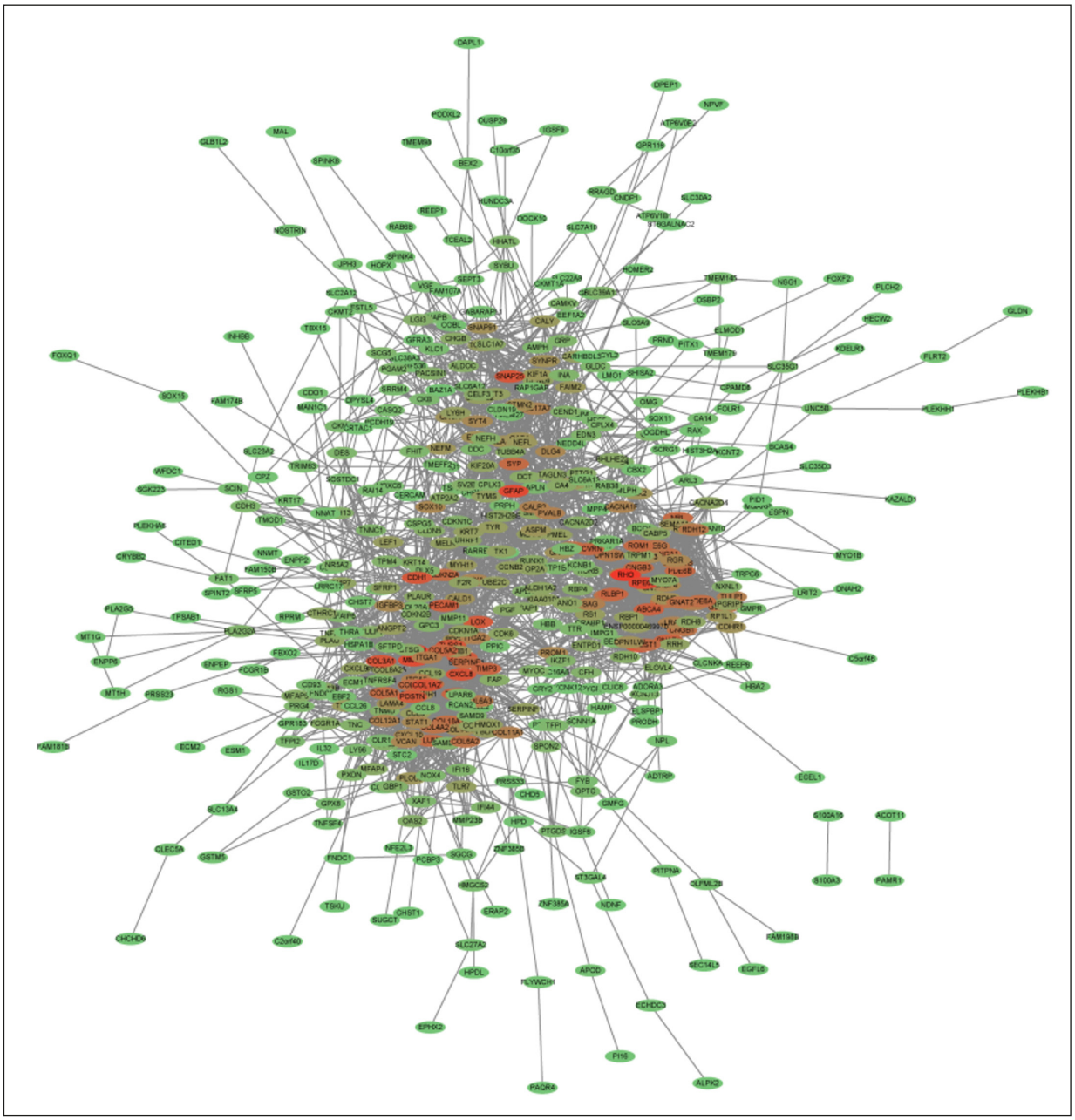

Figure 4. The PPI network of DEGs.

It is one of the microvascular complications of diabetes. People are more likely to develop cataracts at an early age and are twice as likely to develop glaucoma as people without diabetes[13]. It is estimated[13] that by 2030, the global number of DR patients will rise to 191 million[14]. In China, the prevalence of DR in patients with diabetes is 20-40%. DR damages the retinal microvascular system, makes capillaries swell and deform, destroys the blood-retina barrier, and occurs exudation, macular edema, and vision loss. If early diagnosis and appropriate preventive measures can be taken, the visual impairment caused by DR can be reduced to a certain extent[15]. When the disease progresses to new blood vessels, their growth distorts the retina's microvascular system, causing retinal detachment and ultimately blindness[15].

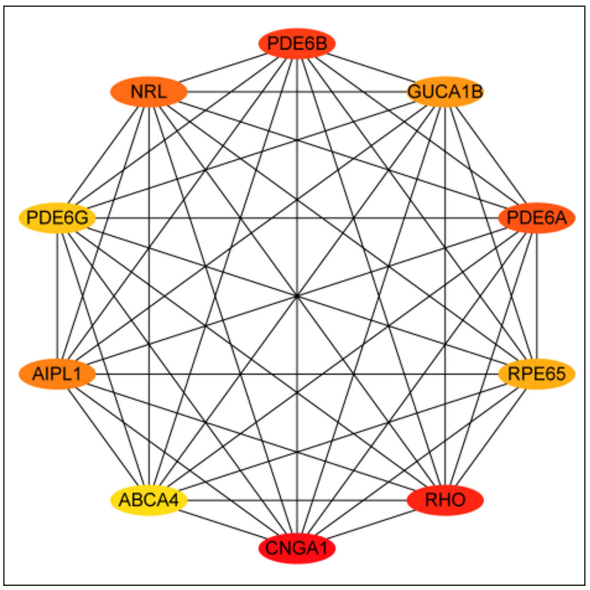

Figure 5. The 10 hub genes of the PPI network.

In recent years, bioinformatics analysis[16-20] has been widely utilized to identify new therapeutic targets and diagnostic markers for various diseases, including cancer, diabetes, hypertension, and other diseases. In the present study, we identified 592 DEGs, including 203 upregulated and 389 downregulated, in the GSE60436 dataset. Furthermore, GO analysis indicated that the DEGs were mostly involved in visual perception, phototransduction, detection of visible light, detection of light stimuli, detection of abiotic stimuli, detection of external stimuli, sensory system development, extracellular structure organization, extracellular matrix organization and sensory perception of light stimuli at the level of BP. In addition, KEGG pathway enrichment analysis showed that the DEGs were enriched significantly within PI3K-Akt signaling pathway, human papillomavirus infection, ECM-receptor interaction, focal adhesion and protein digestion and absorption. There was a limitation we should point out and address in the future. The present data should be confirmed by experimental data such as gene expression or polymorphism[21].

Conclusions

In conclusion, *CNGA1*, *PDE6G*, *RHO*, *ABCA4*, *PDE6A*, *PDE6B*, *NRL*, *RPE65*, *GUCA1B*, and *AIPL1* may be potential biomarkers and therapeutic targets for DR.

Availability of Data and Materials

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

Acknowledgments

We sincerely acknowledge some students from Jiangxi University of Chinese Medicine.

Authors' Contributions

Xiong Jianghao, Cheng Shaomin conceived and designed the manuscript. Xiong Jianghao, Chen Jiali, Liang Jiayu, Zhang Fenfen, and Cheng Shaomin wrote the manuscript. Xiong Jianghao, Chen Jiali, Liang Jiayu, and Zhang Fenfen collected and analyzed the references. Xiong Jianghao, Chen Jiali, and Cheng Shaomin checked, proofread, and polished the manuscript.

Funding

This study was supported in part by the Research project for young and middle-aged backbone talents of Traditional Chinese Medicine in Jiangxi Province, the Key Laboratory of TCM Syndrome Foundation of Jiangxi Administration of Traditional Chinese Medicine and the Science and Technology Innovation Team Development Program of Jiangxi University of Chinese Medicine (No. CXTD22016).

Conflict of Interests

The authors declare that they have no competing interests.

Ethics Approval

Patient data were collected retrospectively from electronic health records. The present study was approved by the Ethics Committee of Jiangxi University of Chinese Medicine and was in accordance with the 1975 Declaration of Helsinki.

Informed Consent

Not applicable.

ORCID ID

Shaomin Cheng: 0000-0002-8737-8916

References

1) Saeedi P, Petersohn I, Salpea P, Malanda B, Karuranga S, Unwin N, Colagiuri S, Guariguata L, Motala AA, Ogurtsova K, Shaw JE, Bright D, Williams R; IDF Diabetes Atlas Committee. Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition. Diabetes Res Clin Pract 2019; 157: 107843.

2) Chen X, Qin L, Li P, Mo W. Cyfip1 is downregulated in acute lymphoblastic leukemia and may be a potential biomarker in acute lymphoblastic leukemia. Tumour Biol 2016; 37: 9285-9288.
3) Chen X, Wang Z, Yan Y, Li P, Yang Z, Qin L, Mo W. XRCC3 C18067T polymorphism contributes a decreased risk to both basal cell carcinoma and squamous cell carcinoma: evidence from a meta-analysis. PloS One 2014; 9: e84195.
4) Jin X, Yin S, Zhang Y, Chen X. Association between TLR2 + 2477G/A polymorphism and bacterial meningitis: a meta-analysis. Epidemiology and infection 2018; 146:642-647.
5) Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Muertter RN, Holko M, Ayanbule O, Yefanov A, Soboleva A. NCBI GEO: archive for functional genomics data sets--10 years on. Nucleic Acids Res 2011; 39: D1005-1010.
6) Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, Lee H, Zhang N, Robertson CL, Serova N, Davis S, Soboleva A. NCBI GEO: archive for functional genomics data sets--update. Nucleic Acids Res 2013; 41: D991-995.
7) Ishikawa K, Yoshida S, Kobayashi Y, Zhou Y, Nakama T, Nakao S, Sassa Y, Oshima Y, Niiro H, Akashi K, Kono T, Ishibashi T. Microarray analysis of gene expression in fibrovascular membranes excised from patients with proliferative diabetic retinopathy. Invest Ophthalmol Vis Sci 2015; 56: 932-946.
8) Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X, Yu G. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. Innovation (Camb) 2021; 2: 100141.
9) Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res 2019; 47: D607-D613.
10) Smoot ME, Ono K, Ruscheinski J, Wang PL, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. Bioinformatics 2011; 27: 431-432.
11) Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst Biol 2014; 8: S11.
12) Chen X, Li P, Yang Z, Mo WN. Expression of fragile histidine triad (FHIT) and WW-domain oxidoreductase gene (WWOX) in nasopharyngeal carcinoma. Asian Pac J Cancer Prev: APJCP 2013; 14: 165-171.
13) Sharma S, Oliver-Fernandez A, Liu W, Buchholz P, Walt J. The impact of diabetic retinopathy on health-related quality of life. Curr Opin Ophthalmol 2005; 16: 155-159.
14) Jin X, Yin S, Zhang Y, Chen X. Quantitative assessment of the association between IL-10 -592 A/C polymorphism and Kawasaki disease risk in Chinese population: evidence from a meta-analysis. Cardiol Young 2018; 28: 811-815.
15) Calderon GD, Juarez OH, Hernandez GE, Punzo SM, De la Cruz ZD. Oxidative stress and diabetic retinopathy: development and treatment. Eye (Lond) 2017; 31: 1122-1130.
16) Qiu Y, Li H, Xie J, Qiao X, Wu J. Identification of ABCC5 Among ATP-Binding Cassette Transporter Family as a New Biomarker for Hepatocellular Carcinoma Based on Bioinformatics Analysis. Int J Gen Med 2021; 14: 7235-7246.
17) Xie J, Li H, Chen L, Cao Y, Hu Y, Zhu Z, Wang M, Shi J. A Novel Pyroptosis-Related lncRNA Signature for Predicting the Prognosis of Skin Cutaneous Melanoma. Int J Gen Med 2021; 14: 6517-6527.
18) Qiu Y, Li H, Zhang Q, Qiao X, Wu J. Ferroptosis-Related Long Noncoding RNAs as Prognostic Marker for Colon Adenocarcinoma. Appl Bionics Biomech 2022; 2022: 5220368.
19) Xie J, Chen L, Sun Q, Li H, Wei W, Wu D, Hu Y, Zhu Z, Shi J, Wang M. An immune subtype-related prognostic signature of hepatocellular carcinoma based on single-cell sequencing analysis. Aging (Albany NY) 2022; 14: 3276-3292.
20) Li C, Qu L, Matz AJ, Murphy PA, Liu Y, Manichaikul AW, Aguiar D, Rich SS, Herrington DM, Vu D, Johnson WC, Rotter JI, Post WS, Vella AT, Rodriguez-Oquendo A, Zhou B. AtheroSpectrum Reveals Novel Macrophage Foam Cell Gene Signatures Associated With Atherosclerotic Cardiovascular Disease Risk. Circulation 2022; 145: 206-218.
21) Zhou X, Cheng FC, Wang HL. Correlations of AQP4 expression and polymorphism with diabetic retinopathy. Eur Rev Med Pharmacol Sci 2021; 25: 1169-1176.