82156_Auto_Edited.docx

WORD COUNT	3008	TIME SUBMITTED	03-FEB-2023 05:18PM
		PAPER ID	96223928

Name of Journal: *World Journal of Clinical Cases*
Manuscript NO: 82156
Manuscript Type: ORIGINAL ARTICLE

Clinical and Translational Research

Analysis of differentially expressed genes related to cerebral ischaemia in young rats based on the GEO database

Differentially expressed genes of cerebral ischaemia

yu xia, han liu, rui zhu

Abstract

BACKGROUND

The incidence rate of cerebral infarction in young people was increasing day by day, the age of onset tended to be younger, and its internal pathogenesis and mechanism were very complicated, which led to greater difficulties in treatment. Therefore, it was essential to analyze the key pathway that affected the onset of young cerebral infarction from the perspective of genetics.

AIM

The aim of this study was to compare the differentially expressed genes in the brain tissue of young and aged rats with middle cerebral artery occlusion and to analyse their effect on the key signalling pathway of cerebral ischaemia in young rats.

METHODS

The GEO 2R online analysis tool was used to analyse the differentially expressed genes in the GSE166162 gene chip regarding the incidence of cerebral ischaemia in young and aged groups of rats. DAVID 6.8 software was further used to filter the differentially expressed genes. These genes were subjected to GO function analysis and KEGG pathway enrichment analysis to determine the key gene pathway that affected the occurrence of cerebral ischaemia in young rats.

RESULTS

Thirty-five differentially expressed genes (such as Igf2, Col1a2, Sfrp1, *etc.*) were obtained; 73 GO enrichment analysis pathways were mainly involved in biological processes such as drug response, amino acid stimulation response, blood vessel development, various signalling pathways, and enzyme regulation. They involved molecular functions such as drug binding, protein binding, dopamine binding, metal ion binding, and dopamine neurotransmitter receptor activity. KEGG pathway enrichment analysis showed a significantly enriched pathway: the c-AMP signalling pathway.

CONCLUSION

The c-AMP signalling pathway might be the key pathway in the intervention of young cerebral infarction.

Key Words: GEO database; Young cerebral infarction; Rats; Differential gene enrichment analysis

xia Y, liu H, zhu R. Analysis of differentially expressed genes related to cerebral ischaemia in young rats based on the GEO database. *World J Clin Cases* 2023; In press

Core Tip: This article mainly uses the analysis technology related to biogenetic informatics to retrieve the samples of young cerebral ischemia rats from the authoritative global GEO database, find out the differential genes related to the onset of young rats through the differential gene analysis method, and further carry out GO function enrichment analysis and KEGG function enrichment analysis, finally obtain the key gene pathways that affect young ischemic stroke. It provides clues for further realizing the transformation of basic medicine into clinical practical application, so as to finally realize the precise target intervention of young people with ischemic stroke and reduce the disability rate and mortality rate.

INTRODUCTION

Cerebral infarction has the characteristics of a high disability rate, high mortality rate and high recurrence rate. It has been the second leading cause of death in the world and seriously endangers human health. With rapid economic and social development, the incidence of cerebral infarction in young people is gradually increasing, and the age of onset tends to be younger. Compared with cerebral infarction in elderly individuals, the aetiology of cerebral infarction in young people is more diversified, which leads to certain difficulties in diagnosis and treatment.

Studies have shown that there are approximately 2 million new young people with cerebral infarction every year in the world, but its treatment methods and effects are still limited[1]. Young people are the backbone of the family, society, and even the country. Once disabled, they will cause considerable damage to society and family. The GEO (Gene Expression Omnibus) database is a public gene expression database created and maintained by NCBI in the United States in 2000. It contains a large amount of gene expression data, such as next-generation sequencing data, chip sequencing, and single-cell sequencing (including rich human clinical data and various animal model data). To find suitable intervention targets and more reasonable and effective treatment methods, we analysed the gene chip data of young cerebral infarction rats and old infarcted rats in the GEO database to identify the key gene pathways that distinguish these groups from one another.

MATERIALS AND METHODS

Selection of subjects

The GEO database (http://www.ncbi.nlm.nih.gov/geo) was searched with the key word "middle cerebral artery occlusion (MCAO)", the deadline was September 2022 and the gene chip GSE166162 of the brain tissue of rats in the young MCAO group and the aged MCAO group was downloaded. The GSE166162 gene chip included male rats aged 3-4 mo and weighing 340 g-400 g and male rats aged 18-20 mo and weighing 585 g-685 g. In this study, the gene chip data of brain tissue was identified according to the principles of rat grouping, age, sampling site, sex, and strain to avoid the influence of differential genes on other factors. Male SD rats aged 3-4 mo and male SD rats aged 18-20 mo were used as the research subjects, and the brain tissue was used as the sampling site. Finally, three rat samples in the young MCAO group were numbered GSM5065037, GSM5065038, and GSM5065039, and three rat samples in the aged MCAO group were numbered GSM5065040, GSM5065041, and GSM5065042 (Table 1).

Differential gene analysis

The GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r) online analysis tool was used to perform differential gene analysis on the GSE166162 gene chip. With logFC absolute value ≥ 1 and P < 0.05 as screening conditions, the differentially expressed genes in the brain tissue of the young MCAO group rats and the aged MCAO group rats were selected. In addition, up and downregulated genes significantly regulated by young cerebral ischaemia rats were screened out by adjusting P values.

Statistical analysis

To explain the pathways affected by differential genes in the brain tissue of cerebral infarction rats, all the differentially expressed genes were introduced into DAVID 6.8 (https://david.ncifcrf.gov/) for GO functional analysis and KEGG pathway enrichment analysis. The overall analysis idea and specific analysis steps of this paper were shown in a flow chart (Fig. 1).

RESULTS

Outcome of differential gene analysis

Using GEO2R to conduct differential gene analysis on the GSE166162 gene chip, Thirty-five differentially expressed genes were obtained. Among them, twenty-three genes were upregulated with positive logFC (Table 2), and twelve genes with negative logFC were downregulated (Table 3). In addition, the significantly upregulated and downregulated genes were further screened after adjusting the P value with volcano maps (Table 4-5). A volcano map can visually display red upregulated genes and blue downregulated genes (Fig. 2). The above methods were used to discern related genes that could play a key role in the pathogenesis of young infarcted rats. The abnormal expression of these genes might affect biochemical processes, molecular functions and related gene pathways in rats.

Outcome of GO functional enrichment analysis

The DAVID 6.8 online analysis tool was used to perform GO functional enrichment analysis on the above Thirty-five differentially expressed genes, and one hundred

and eleven GO items were obtained. With P<0.05 as a significant enrichment criterion, seventy-three pathways were obtained. Among them, 54 (48.65%) were enriched in biological processes, including drug response, amino acid stimulation response, vascular development, various signalling pathways, enzyme regulation and other biological processes; 10 (9.01%) were enriched in cell composition, mainly including extracellular matrix, collagen trimer and troponin complex; 9 (8.11%) were enriched in molecular functions, mainly including drug binding, protein binding, dopamine binding, metal ion binding and dopamine neurotransmitter receptor activity (Table 6).

Outcome of KEGG pathway enrichment analysis

Using the DAVID 6.8 online analysis tool, the KEGG pathway enrichment analysis was performed on the thirty-five differentially expressed genes, and a total of seven pathways were enriched. With P<0.01 as the significant enrichment criterion, a significant enrichment pathway was obtained. This pathway was the c-AMP signalling pathway (Table 7).

DISCUSSION

Cerebral infarction is an important cause of disability and death worldwide. Although traditionally considered a disease of elderly individuals, infarction in young people is increasingly a global public health problem, and the average age of stroke onset is gradually decreasing. The incidence of infarction and hospitalization rates is increasing among young people in some countries[1]. The Infarction in young people is generally considered to be different from that in older patients in terms of genetic pathogenesis, and its onset is a complex process of multigene and multipathway interactions, which is an important reason for the difficulty in treating the disease [2, 3]. Therefore, revealing the pathogenesis of young ischaemic infarction at the molecular level has become a research hotspot in the treatment of this disease. In this study, suitable young rats with ischaemic infarction were selected from the GEO database as the research subjects, and biological gene chip technology was used

to identify more important genetic pathways to intervene and reduce the incidence of young ischaemic infarction to the greatest extent.

Studies have found that the collagen type I alpha 2 (COL1A2) gene, which is a susceptibility gene of collagen-related diseases such as systemic sclerosis, contains a specific combination of two dinucleotide repeats and is involved in the regulation of gene expression. The COL1A2 gene overaccumulates in affected organs or tissues, eventually leading to a systemic inflammatory immune response that causes damage to blood vessels in young infarcts [4, 5]. The above findings suggest that COL1A2 may be closely related to ischaemic cerebrovascular diseases, such as vascular inflammation caused by systemic immune system abnormalities. This study is the first to find that COL1A2 is associated with ischaemic stroke in young adults, which suggests that COL1A2 could be modulated to reduce vascular inflammation and stabilize the vascular intima, thereby avoiding thrombus aggregation caused by vascular intimal injury. From the perspective of the c-AMP pathway, the increase in the c-AMP concentration could inhibit the migration and intimal hyperplasia of vascular smooth muscle cells (VSMCs), promote the degradation of type I collagen by lysosomes, and reduce the content of type I collagen secreted into and out of VSMCs to protect them [6]. The report is for further clinical research. Collagen type VIII alpha 2 (COL8A2) is a protein-coding gene that encodes the α2 chain of type VIII collagen, a component of the vascular endothelium, which is responsible for the migration of VSMCs. It is necessary for proliferation and thus plays an important role in maintaining vessel wall integrity and structure [7]. This indicates that COL8A2 is similar to COL1A2, and regulation of COL8A2 can also reduce the incidence of stroke in young adults. It has been reported in the past that tumour necrosis factor-α (TNF-α) is an important proinflammatory cytokine secreted by vascular endothelial cells that mediates the production of many inflammatory mediators. Increased TNF-α is an independent risk factor for acute cerebral infarction [8], which is consistent with the conclusion of this study. The systemic inflammatory immune response may play a very important role in the pathogenesis of stroke in young people, and it is worthy of further exploration.

Previous studies have confirmed that the secreted frizzled related protein 1 (SFRP1) gene is a genetic predisposition gene for large atherosclerotic cerebral infarction [9], which is consistent with the results of this study. Therefore, young people with high risk for ischaemic stroke (such as long-term hypertension, heavy smoking, severe abdominal obesity, *etc.*) are genetically tested to screen out high-risk stroke patients who carry this susceptibility gene to minimize stroke morbidity. Studies have shown that insulin-like growth factor 2 (IGF2) is abundantly expressed in the central nervous system, and its deficiency is closely related to neuropsychiatric diseases such as vascular dementia [10]. This suggests that severe deficiency of IGF2 may cause the repeated occurrence of strokes in such patients, which in turn leads to a stepwise aggravation of the condition of patients with vascular cognitive impairment and eventually leads to the occurrence of diseases such as vascular dementia. This is also a scientific proposition worthy of in-depth exploration. Studies have shown that elevated serum troponin T is a predictor of poor prognosis in acute stroke, and elevated troponin T concentrations are associated with an increased risk of death in hospitalized patients [11], which is consistent with the findings of this study. Therefore, in addition to paying attention to the common haematological risk factors for stroke in young people, clinicians also need to pay attention to troponin T, an indicator of cardiac injury. Preleukaemia transcription factor 3 (PBX3) and ATP-binding cassette transporter subfamily A member 8 (ABCA8) have been reported in malignant tumours. They play a key role in the pathogenesis of cancer, and high levels of PBX3 and ABCA8 are closely associated with poor prognosis in cancer patients. Their association with tumorigenesis indirectly suggests that they may contribute to a hypercoagulable state in normal young patients, leading to thrombosis and eventual stroke [12, 13]. This indicates that genes regulating tumorigenesis may be involved in the pathophysiological process of cerebral infarction, which is worthy of further exploration. Quinoid dihydrobiopterin is reduced to tetrahydrobiopterin under the action of quinoid dihydropteridine reductase (QDPR). Tetrahydrobiopterin can affect the vascular endothelium in various ways to promote the occurrence and development of

cardiovascular and cerebrovascular diseases. For example, tetrahydrobiopterin deficiency can promote atherosclerosis and the production of oxygen free radicals [14]. This suggests that QDPR may play an important role in the pathogenesis of young infarcts. Based on the above findings, we speculate that atherosclerosis may also play an important role in the incidence of stroke in young adults, especially those after the age of thirty-five with poor diet and living habits (because they may have a higher proportion of diseases such as hypertension).

In addition, we obtained seventy-three significantly enriched pathways after GO functional enrichment analysis, of which fifty-four were significantly enriched in biological processes including drug reactions, various signalling pathways, enzyme regulation and other biological processes; ten were significantly enriched with a focus on cell composition, mainly including extracellular matrix, collagen trimer, troponin complex, *etc.*; nine were significantly enriched in molecular function, mainly including drug binding, protein binding, dopamine binding and dopamine neurotransmitter receptors. Finally, KEGG pathway enrichment analysis was performed on thirty-five differentially expressed genes to obtain a significantly enriched pathway: the cyclic adenosine monophosphate (c-AMP) signalling pathway. Studies have shown that the main target of the c-AMP was the c-AMP-dependent protein kinase A (PKA) and exchange protein directly activated by the c-AMP (EPAC). We might reduce the incidence of ischaemia in young rats by interfering with corresponding targets, thereby reducing their disability and mortality. The c-AMP signalling pathway can promote vascular endothelial repair and inhibit vascular intimal proliferation and platelet aggregation by regulating downstream effector molecules, thereby inhibiting vascular remodelling. In addition, activation of the c-AMP pathway can also alleviate ischaemic stroke-reperfusion injury, promote neuronal repair, and improve poststroke cognitive function [15, 16]. The c-AMP signalling pathway may be a key pathway for the intervention of stroke-related diseases such as young cerebral infarction and vascular cognitive impairment and has extremely important clinical significance and research value.

To date, there are no relevant research paper at home and abroad reporting the key signalling pathway of ischemic stroke in young rats. In this paper, we obtained that the c-AMP signalling pathway is the key signalling pathway for young rats developing acute ischemic infarction through the methods of related gene analysis, which has extremely significant clinical significance. However, the subjects of this study are rats, and the sample size is relatively small. If more gene expression profile data are published in the future, further in-depth analysis could be conducted to verify the research results. Finally, we need to collect a certain number of clinical cases in the future to verify whether the research results are suitable for human beings, so as to realize the clinical transformation of major basic scientific discoveries.

CONCLUSION

Our study demonstrated that the c-AMP signalling pathway may be the key pathway in the intervention of young cerebral infarction. Further fundamental studies are needed to comprehensively explore the impact of the key pathway response on young infarction to finally achieve clinical transformation.

ARTICLE HIGHLIGHTS

Research background

The incidence rate of cerebral infarction in young people was increasing day by day, the age of onset tended to be younger, and its internal pathogenesis and mechanism were very complicated, which led to greater difficulties in treatment. Therefore, it was essential to analyze the key pathway that affected the onset of young cerebral infarction from the perspective of genetics.

Research motivation

The aim of this study was to compare the differentially expressed genes in the brain tissue of young and aged rats with middle cerebral artery occlusion and to analyse their effect on the key signalling pathway of cerebral ischaemia in young rats.

Research objectives

The aim of this study was to compare the differentially expressed genes in the brain tissue of young and aged rats with middle cerebral artery occlusion and to analyse their effect on the key signalling pathway of cerebral ischaemia in young rats.

Research methods

The GEO 2R online analysis tool was used to analyse the differentially expressed genes in the GSE166162 gene chip regarding the incidence of cerebral ischaemia in young and aged groups of rats. DAVID 6.8 software was further used to filter the differentially expressed genes. These genes were subjected to GO function analysis and KEGG pathway enrichment analysis to determine the key gene pathway that affected the occurrence of cerebral ischaemia in young rats.

Research results

Thirty-five differentially expressed genes (such as Igf2, Col1a2, Sfrp1, *etc.*) were obtained; 73 GO enrichment analysis pathways were mainly involved in biological processes such as drug response, amino acid stimulation response, blood vessel development, various signalling pathways, and enzyme regulation. They involved molecular functions such as drug binding, protein binding, dopamine binding, metal ion binding, and dopamine neurotransmitter receptor activity. KEGG pathway enrichment analysis showed a significantly enriched pathway: the c-AMP signalling pathway.

Research conclusions

The c-AMP signalling pathway might be the key pathway in the intervention of young cerebral infarction.

Research perspectives

Young cerebral infarction

ACKNOWLEDGEMENTS

This work was supported by the Anhui Medical University Research Fund and hospital level scientific research project of the Third People's Hospital of Hefei.

82156_Auto_Edited.docx

ORIGINALITY REPORT

4%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet	41 words — 1%
2	pesquisa.bvsalud.org Internet	17 words — 1%
3	repository.bilkent.edu.tr Internet	14 words — < 1%
4	www.researchgate.net Internet	14 words — < 1%
5	www.science.gov Internet	13 words — < 1%
6	Lan Hainan, Liu Huilin, Muhammad Akram Khan, Zheng Xin, Yang YuJiang, Zhang Hui, Yao Naiquan. "The basic route of the nuclear translocation porcine growth hormone (GH)-growth hormone receptor (GHR) complex (pGH/GHR) in porcine hepatocytes", General and Comparative Endocrinology, 2018 Crossref	12 words — < 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 12 WORDS
EXCLUDE BIBLIOGRAPHY	OFF	EXCLUDE MATCHES	< 12 WORDS