

doi: 10.3978/j.issn.2095-6959.2020.09.017

View this article at: <http://dx.doi.org/10.3978/j.issn.2095-6959.2020.09.017>

急性缺血性卒中相关基因的筛选和验证

蒋蔚¹, 万蛟龙¹, 包元飞², 王新³

(1. 南京医科大学附属无锡市第二人民医院脑科中心, 江苏 无锡 214000; 2. 南通大学附属南通第三人民医院神经内科, 江苏 南通 226001; 3. 江苏省神经再生重点实验室, 江苏 南通 226001)

[摘要] **目的:** 通过分析基因表达谱, 探讨缺血性脑卒中(ischemic stroke, IS)的分子机制和诊断标志物。**方法:** 下载GEO数据库中缺血性卒中患者和对照组的基因芯片数据, 使用Pubmed提供的GEO2R软件和R语言筛选出具有显著意义的差异基因(differential genes, DEGs); 利用Metascape进行基因本体(gene ontology, GO)与酶和通路数据库(Reactome)基因集分析; 使用STRING在线分析工具和Cytoscape软件分析出关键基因; 最后, 收集急性卒中患者外周血, 通过RT-PCR进行验证。**结果:** 通过分析IS患者外周血基因芯片数据集, 共筛选出20个差异表达基因; 基因注释和信号途径分析表明, 这些差异表达基因主要参与中性粒细胞脱颗粒、对外来生物刺激反应、细胞因子生产、有机阴离子运输、对脂肪酸的反应和中性粒细胞迁移等。使用Cytoscape和MCODE插件进一步分析, 筛选出5个关键基因(MMP-9, OLFM4, ORM1, FOLR3和ARG1); RT-PCR证实ARG1和MMP9 mRNA在IS患者外周血中过表达。**结论:** 缺血性卒中伴随多个基因表达异常, 本研究为探索IS的发病机制和开发新的诊疗策略提供了线索。

[关键词] 缺血性卒中; 精氨酸酶-1; 基质金属蛋白酶-9

Screening and verification of genes related to ischemic stroke

JIANG Wei¹, WAN Jiaolong¹, BAO Yuanfei², WANG Xin³

(1. Brain Center, Wuxi Second People's Hospital, Nanjing Medical University, Wuxi Jiangsu 214000; 2. Department of Neurology, Nantong Third People's Hospital, Nantong University, Nantong Jiangsu 226001; 3. Jiangsu Key Laboratory of Nerve Regeneration, Nantong Jiangsu 226001, China)

Abstract **Objective:** The objective of this study was to provide insight into the molecular mechanisms of acute ischemic stroke (IS) through gene expression profiling analysis. **Methods:** The gene chip data of ischemic stroke patients and the control group in the GEO database were downloaded. Differential genes were analyzed by GEO2R software and the R language. Metascape was used to analyze gene ontology (GO) and gene sets of enzyme and pathway analysis (Reactome). The hub genes were identified using the STRING online analysis tool and Cytoscape software. Finally, quantitative real-time polymerase chain reaction (qRT-PCR) was performed to verify

收稿日期 (Date of reception): 2020-03-16

通信作者 (Corresponding author): 蒋蔚, Email: jianghaot1986@163.com

基金项目 (Foundation item): 南通市自然科学基金 (MS12017021-4)。This work was supported by the Natural Science Foundation of Nantong, Jiangsu Province, China (MS12017021-4).

the expression of DEGs in the peripheral blood of IS patients. **Results:** A 20-gene profile was identified in the whole blood of ischemic stroke patients from the gene expression datasets. Gene annotation and signal pathway analysis showed that these differentially expressed genes were mainly involved in the response of neutrophils to extraneous biological stimulation, the response of cytokines to organic anion production, transport to fatty acids, and the migration of neutrophils, etc. Five genes (mmp-9, OLFM4, ORM1, FOLR3, and ARG1) were identified as hub gene by Cytoscape and MCODE software. Finally, ARG1 and MMP9 mRNA overexpressed in peripheral blood of IS patients compared with normal control confirmed by RT-PCR. **Conclusion:** The dysregulation of several genes may involve the process of ischemic stroke (IS). Our finding provided clues for exploring mechanisms and developing novel diagnostic and therapeutic strategies for IS.

Keywords ischemic stroke; ARG1; MMP9

全球范围内, 卒中是第二大死亡原因, 随着年龄的增长, 卒中患病率在男性和女性中都有所增加^[1]。缺血性脑卒中(ischemic stroke, IS)是由于脑内动脉狭窄或闭塞, 引起脑供血和供氧不足, 从而导致肢体运动和感觉障碍等症状的脑血管疾病, 是脑卒中的主要类型, 约占脑卒中病例的87%^[2-3]。IS病理生理进程中多个表达异常的基因, 对IS的发生、发展和转归起重要作用。

本文通过分析急性IS后外周全血的基因表达谱, 筛选并验证急性缺血性脑卒中发病过程中的起重要调控作用的基因, 以深入理解IS的分子机制。

1 资料与方法

1.1 基因芯片数据

从美国国家生物技术信息中心的 Gene Expression Omnibus(GEO)数据库(<https://www.ncbi.nlm.nih.gov/geo/>)下载IS研究数据集(GSE16561)。该数据集系缺血性脑卒中患者和正常人外周血标本的mRNA芯片数据^[4]。

1.2 分析差异表达基因

采用GEO2R和R studio软件对下载的基因表达谱数据进行分析, 再采用非配对t检验分析差异基因(differential genes, DEGs), 筛选标准需同时满足以下条件: 1) $P < 0.05$; 2) 组间差异表达倍数(fold change) ≥ 2.0 。

1.3 基因功能注释和信号通路分析

使用Metascape(<http://Metascape.org>)进行基因

本体(GO)富集分析和酶和通路数据库(Reactome)基因集分析, 以 $P < 0.05$ 为有显著意义的基因功能和信号通路。

1.4 蛋白-蛋白相互作用分析(PPI)

使用STRING(<https://string-db.org/>)在线分析工具, 构建PPI网络, 并将分析结果导入Cytoscape软件进行可视化, 使用MOCDE插件分离PPI网络中的关键子网络。

1.5 RT-PCR 验证 DEGs 在卒中患者血液中的表达

验证的血液样系无锡市第二人民医院神经内科收治的45例IS患者和15例正常人。IS患者年龄范围60~83岁, 中位年龄67岁, 其中28例男性, 17例女性, 均经过MRI诊断, 排除其他神经系统疾病、心脏栓塞、短暂性脑缺血发作、出血性梗死、隐性脑血管畸形、外伤性脑血管病的患者, 本次研究对象均为溶栓适应证以外就诊患者, 血液样本收集时间为症状发作后5~20 h。15例正常对照组血液样本, 来自无锡市第二人民医院体检中心, 年龄>60周岁, 其中8例男性, 7例女性, 均无脑卒中史, 无脑外伤及手术史, 无心脏手术史, 无神经系统疾病史。使用Paxgene Blood RNA Extraction Kit (Qiagen)提取和纯化全血中总RNA, 经反转录后获得cDNA; 以GAPDH(NM_002046.7)作为内参基因, 利用FastStart Universal SYBR Green Master(Roche)试剂盒, qPCR检测IS患者组和正常对照组中ARG1(NM_000045.4), FOLR3(NM_000804.4), MMP9(NM_004994.3), OLFM4(NM_006418.5)和ORM1(NM_000607.3)表达。利用Primer 3在线引物设计工具(<http://primer3.ut.ee>)设计各基因qPCR引物, 如表1所示。

表1 RT-PCR引物列表

Table 1 RT-PCR primer list

基因名	正向引物(5'-3')	反向引物(5'-3')
<i>GAPDH</i>	CCAGAACATCATCCCTGCCT	CCTGCTTCACCACCTTCTTG
<i>ARG1</i>	GTGGAAGAAGGCCCTACAGT	GCTTTCACACAGACCTTGG
<i>FOLR3</i>	CTTCTGGCTTTGGTGACTGC	GCAGGCATTCCTTCCAGG
<i>MMP9</i>	TCTCCCTGGAGACCTGAGA	CTATCCAGCTCACCGGTCTC
<i>OLFM4</i>	CCGGCTCTCATTCTCTCTA	AGAACCTCCGCTGCCTAAG
<i>ORM1</i>	GAGAGTACCAGACCCGACAG	TCAGCATAGACAGACAGCCC

2 结果

2.1 数据集样本的临床信息

GSE16561数据集共63例样本, 39例急性缺血性脑血管综合征(acute ischemic cerebrovascular syndrome, AICS)均经过MRI诊断分型(TOAST分型), 17例为心源性脑栓塞, 11例为不明原因型, 5例为大动脉栓子/血栓形成, 6例为小血管或其他原因; 24例对照组均为成年人。从症状发作到抽血的时间约(10±6.5) h。两组之间的性别没有差异; 但是, 卒中组患者年龄较大($P<0.001$, 表2), 并且正如预期的那样, 卒中组的血管危险因素更为常见。

2.2 急性卒中患者血液中的差异表达基因 (DEGs)

从GEO下载IS患者血液基因检测数据集(GSE16561)。与正常对照组相比, 以 $\text{Log}_2\text{FC} \geq 1$ (差异倍数 ≥ 2.0)和校正后 $P<0.05$ 为限定条件筛选DEGs, 并使用火山图(图1)和热图(图2)将结果可视化, 鉴定出20个差异表达基因, 其中19个表达上调, 1个表达下调(表3)。

2.3 DEGs 的富集分析

使用功能富集分析初步理解DEGs的生物学意

义, 将DEGs输入Metascape(<http://Metascape.org>)在线工具进行基因本体(GO)富集分析和酶和通路数据库(Reactome)基因集分析, 以 $P<0.05$ 为有显著意义的基因功能和信号通路。差异基因主要富集在: 中性粒细胞脱颗粒、对外来生物刺激反应、细胞因子生产、有机阴离子运输、对脂肪酸的反应和中心粒细胞迁移(图3)。

2.4 PPI 网络构建与模块分析

使用String在线网络分析工具和Cytoscape软件构建PPI网络图, 将20个DEGs输入String, 共得到12个节点(图4A), 及18条相互关系连线; 使用Cytoscape软件和MOCDE程序分离出主要子网络, 该网络有5个基因(图4B): *ARG1*, *OLFM4*, *MMP9*, *FOLR3*和*ORM1*为DEGs的关键基因。

2.5 RT-PCR 验证关键 DEGs 的表达

基于我们对GEO卒中患者血液的微阵列分析, 选择5种DEGs(*ARG1*, *OLFM4*, *FOLR3*, *MMP9*和*ORM1*), 使用定量聚合酶链反应(reverse transcription-PCR, RT-PCR)进行验证(图5)。与正常对照组相比, RT-PCR证实*ARG1*和*MMP9*表达上调, 差异有统计学意义($P<0.01$)。

表2 卒中组和对照组样本的临床参数分析

Table 2 Analysis of clinical parameters between stroke patients and control subjects

组别	性别(男/女)	年龄/岁	高血压/例	糖尿病/例	高血脂/例	房颤病史/例	缺血性中风史/例	卒中家族史/例
AICS组	17/22	73.1 ± 14.0	25	11	18	6	6	15
对照组	10/14	59.9 ± 9.73	7	2	0	0	2	4
<i>P</i>	0.883	<0.001	0.010	0.068	<0.001	0.048	0.414	0.169

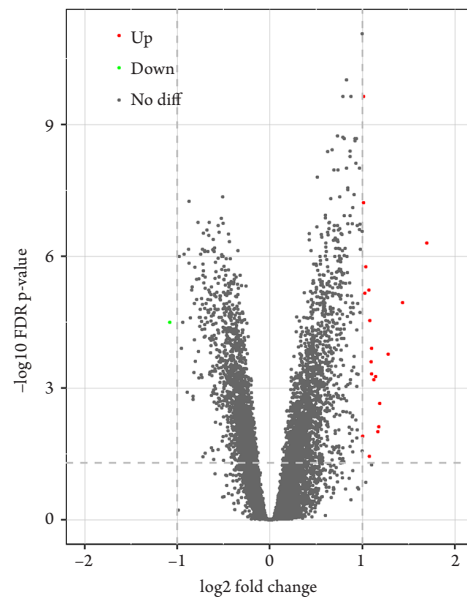


图1 急性IS患者血液中差异表达基因的火山图

Figure 1 Volcanic map of differential expression genes in the blood of IS patients

X轴: 差异倍数; Y轴: P值, 红点: 上调基因; 绿点: 下调基因; 黑点: $|\log_2 FC| < 1$ 或 $P \geq 0.05$ 。

X-axis: multiple of difference; Y-axis: P value. Red point: up-regulated gene; Green Point: down regulated gene; Black point: $|\log_2 FC| < 1$ or $P \geq 0.05$.

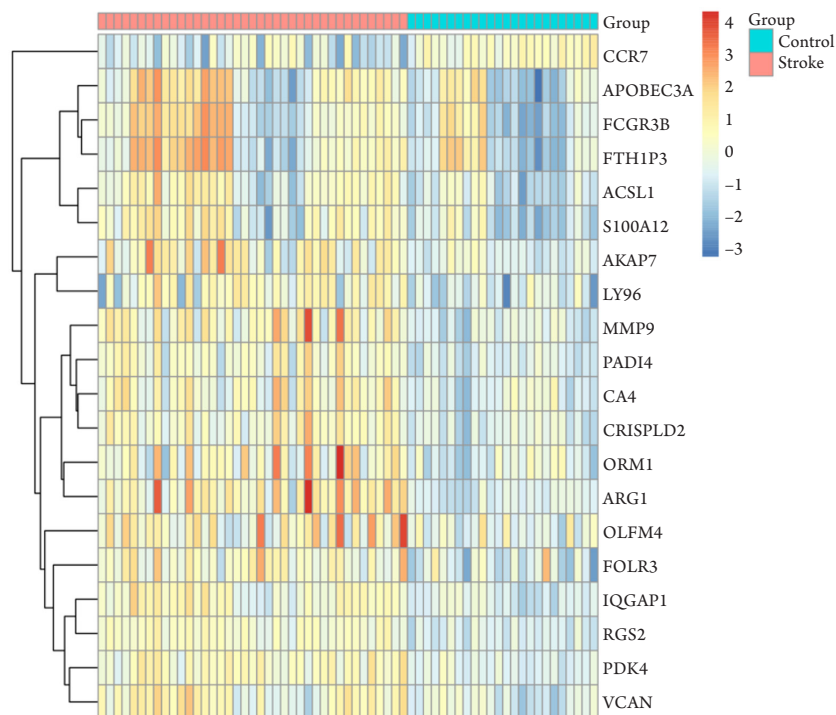


图2 急性IS患者血液中差异表达基因(DEGs)的热图

Figure 2 Heat map of differentially expressed genes (DEGs) in the blood of IS patients

DEGs聚类分析的树状图, 每一列代表一个样本, 淡蓝色: 对照组, 粉红色: 急性卒中患者。每行代表一个基因, 红色和黄色: 相对高表达; 灰蓝色: 相对低表达。

Cluster analysis of the DEGs, each column represents a sample; light blue: control group; pink: acute stroke patients. Each line represents one gene; red and yellow: relatively high expression; gray-blue: relatively low expression.

表3 卒中基因简况

Table 3 Stroke gene profile

Gene	adj.P.Val	Regulation	<i>t</i>	logFC
<i>RGS2</i>	2.29E-10	Up	9.59	1.01
<i>PKD4</i>	6.03E-08	Up	7.72	1.01
<i>ARG1</i>	0.00000489	Up	6.97	1.69
<i>IQGAP1</i>	0.00000172	Up	6.55	1.03
<i>CRISPLD2</i>	0.00000576	Up	6.13	1.07
<i>PADI4</i>	0.0000069	Up	6.06	1.02
<i>MMP9</i>	0.0000112	Up	5.90	1.43
<i>VCAN</i>	0.0000288	Up	5.57	1.08
<i>CCR7</i>	0.0000319	Down	-5.53	-1.08
<i>CA4</i>	0.000122	Up	5.07	1.10
<i>S100A12</i>	0.000167	Up	4.97	1.28
<i>ACSL1</i>	0.000251	Up	4.83	1.09
<i>FOLR3</i>	0.000469	Up	4.61	1.09
<i>AKAP7</i>	0.000541	Up	4.56	1.14
<i>LY96</i>	0.000643	Up	4.50	1.12
<i>ORM1</i>	0.0022	Up	4.05	1.18
<i>FCGR3B</i>	0.00741	Up	3.58	1.17
<i>APOBEC3A</i>	0.00974	Up	3.47	1.16
<i>OLFM4</i>	0.0123	Up	3.38	1.00
<i>FTH1P3</i>	0.0354	Up	2.93	1.07

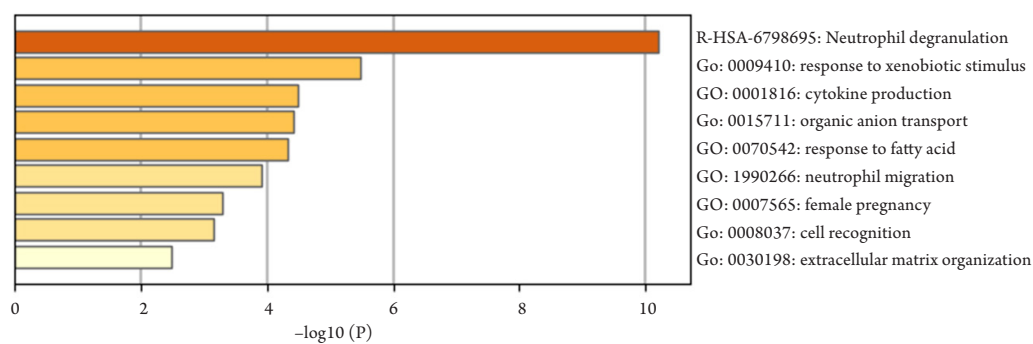


图3 差异基因富集分析

Figure 3 Enrichment analysis of DEGs

X轴表示P值, Y轴表示基因本体注释和信号途径。

The X-axis represents the *P* value, and the Y-axis represents the GO annotation and signal pathway.

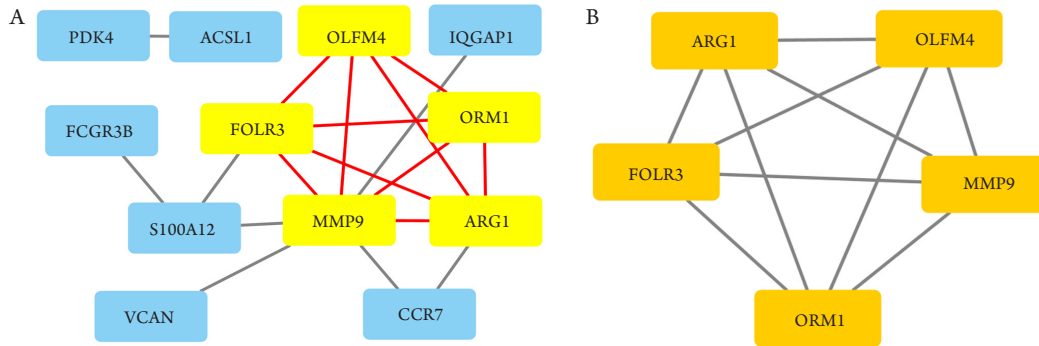


图4 利用STRING在线数据库和Cytoscape软件构建DEGs-PPI网络

Figure 4 Construct the DEGS-PPI network by using the STRING online database and Cytoscape software

(A) DEGs-PPI网络复合体中共有12节点, 节点表示蛋白质, 连线表示蛋白质的相互作用; (B) 利用Cytoscape软件和MCODE应用程序 (degree cutoff=2, node score cutoff=0.2, k-core=2, max.Depth =100) 进行模块分析, 分离出主要子网络。

(A) There are 12 nodes in the complex of the DEGs-PPI network. The nodes represent proteins, and the wires represent protein interactions.

(B) Cytoscape software and MCODE application (degree cutoff =2, node score cutoff =0.2, k-core =2, max-depth =100) were used to separate the main sub-networks.

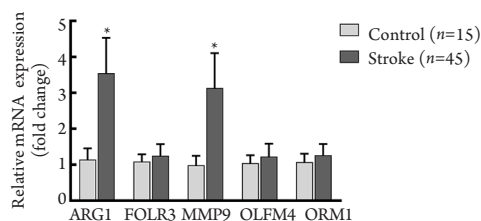


图5 验证IS患者外周血关键DEGs mRNA水平

Figure 5 Expression levels of critical DEGs were verified in patients with IS

X轴表示DEGs, Y轴表示IS和正常对照组之间的倍数变化。Control: 对照组; Stroke: IS组; * $P < 0.05$ 。

The X-axis is the DEGs, and the Y-axis is the multiple changes between IS and the normal control group. Control: Control group; Stroke: IS group; * $P < 0.05$.

3 讨论

在我国, IS发病率持续上升, 是中国居民主要死亡原因之一^[5]。尽管由于神经影像学技术和临床诊疗的进步, 生存率有一定提高, 但发病率和致残率仍然很高。溶栓是IS的最有效治疗方法, 然而, 溶栓剂必须在限定的时间段内施用, IS的早期和快速诊断对于治疗至关重要^[6]。基因表达谱的微阵列芯片可以大规模检查参与生物过程的基因, 反映和预测疾病的病理过程, 被广泛用于研究恶性肿瘤、神经系统和免疫调节紊乱等疾病, 利用基因芯片技术检测IS患者外周血基因表达的变化,

有助于探索IS的发病机制、开发早期诊断和判断预后的生物标志物。

本文通过分析IS患者和正常对照组外周血的基因芯片, 筛选出20个基因在IS(19个表达上调, 1个下调)患者差异表达, 将20个DEGs进行功能注释发现, 对外来生物刺激的反应、细胞因子生产、有机阴离子运输、对脂肪酸的反应和中心粒细胞迁移均为显著的GO项, 中性粒细胞脱颗粒是明显富集的信号途径。进一步分析提示, MMP-9、*OLFM4*、*ORM1*、*FOLR3*和*ARG1*可能是关键基因。RT-PCR验证结果, 证实MMP9和*ARG1* mRNA在IS患者外周血中过表达。侯延伟等^[7]研究报道MMP-9基因rs20544多态性与急性IS患者出血性转化的相关性。Zhong等^[8-9]检测3 186例急性缺血性卒中降压试验参与者外周血的MMP-9水平, 随访3个月后评估预后, 发现缺血性中风急性期血清MMP-9水平升高与认知障碍、死亡和重大残疾风险增加有关, 血清MMP-9可能是IS的重要预后因素。国内尚未检索到*ARG1*在IS的研究报道, 国外Yoo等^[10]研究发现: IS患者外周血中miR-340-5p被下调, 而*ARG1* mRNA被上调, 在人类嗜中性粒细胞和小鼠巨噬细胞细胞系中miR-340-5p的过表达诱导*ARG1*蛋白的下调, 急性卒中通过抑制miR-340-5p上调*ARG1*蛋白的表达。免疫系统平衡反应在IS恢复中起重要作用, 中性粒细胞-淋巴细胞比率(NLR)可能是IS后免疫失调程度的标志。Ashley等^[11]研究报道*ARG1*, NLR和中风严重程度之间存在相关性, 其具体分子机制有待进一步探索, 外周血中*ARG1*

表达在诊断IS和判断预后中的价值有待后续的临床研究验证。

总之, 本研究通过整合分析中获得了新的DEGs和信号通路途径, 提示整合微阵列分析是揭示疾病分子机制的一种好方法; 本研究筛选并验证MMP9和ARG1在IS患者外周血异常表达, 为了解IS的病理、诊断、治疗和预后评估提供了线索。但是, 本研究存在一定的局限性, 用于RT-PCR确认的样本数量很少。为了证实我们的发现并探索关键DEGs在IS中的精确作用, 我们将进行更大样本量的进一步研究。

参考文献

1. Zhou M, Wang H, Zeng X, et al. Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017[J]. *Lancet*, 2019, 394(10204): 1145-1158.
2. Rosamond W, Flegal K, Furie K, et al. Heart disease and stroke statistics--2008 update: a report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee[J]. *Circulation*, 2008, 117(4): e25-e146.
3. Feigin VL, Roth GA, Naghavi M, et al. Global burden of stroke and risk factors in 188 countries, during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013[J]. *Lancet Neurol*, 2016, 15(9): 913-924.
4. Barr TL, Conley Y, Ding J, et al. Genomic biomarkers and cellular pathways of ischemic stroke by RNA gene expression profiling[J]. *Neurology*, 2010, 75(11): 1009-1014.
5. 王陇德, 王金环, 彭斌, 等. 《中国脑卒中防治报告2016》概要[J]. *中国脑血管病杂志*, 2017, 14(4): 217-224.
WANG Longde, WANG Jinghuan, PENG Bin, et al. Summary of "China Stroke Prevention Report 2016"[J]. *Chinese Journal of Cerebrovascular Diseases*, 2017, 14(4): 217-224.
6. Berkowitz A L, Mittal M K, Mclane H C, et al. Worldwide reported use of IV tissue plasminogen activator for acute ischemic stroke[J]. *Int J Stroke*, 2014, 9(3): 349-355.
7. 侯延伟, 郭再玉, 吐尔洪·吐尔逊, 等. MMP-9基因rs20544多态性与急性缺血性卒中患者出血性转化的相关性[J]. *国际脑血管病杂志*, 2019, 27(7): 497-502.
HOU Yanwei, GUO Zaiyu, TUERHONG Tuerxun, et al. Association between MMP-9 gene rs20544 polymorphism and hemorrhagic transformation in patients with acute ischemic stroke[J]. *International Journal of Cerebrovascular Diseases*, 2019, 27(7): 497-502.
8. Zhong C, Yang J, Xu T, et al. Serum matrix metalloproteinase-9 levels and prognosis of acute ischemic stroke[J]. *Neurology*, 2017, 89(8): 805-812.
9. Zhong C, Bu X, Xu T, et al. Serum Matrix Metalloproteinase-9 and Cognitive Impairment After Acute Ischemic Stroke[J]. *J Am Heart Assoc*, 2018, 7(1): e007776.
10. Yoo H, Kim J, Lee AR, et al. Alteration of microRNA 340-5p and arginase-1 expression in peripheral blood cells during acute ischemic stroke[J]. *Mol Neurobiol*, 2019, 56(5): 3211-3221.
11. Petrone AB, O'Connell GC, Regier MD, et al. The role of arginase 1 in post-stroke immunosuppression and ischemic stroke severity[J]. *Transl Stroke Res*, 2016, 7(2): 103-110.

本文引用: 蒋蔚, 万蛟龙, 包元飞, 王新. 急性缺血性卒中相关基因的筛选和验证[J]. *临床与病理杂志*, 2020, 40(9): 2325-2331. doi: 10.3978/j.issn.2095-6959.2020.09.017

Cite this article as: JIANG Wei, WAN Jiaolong, BAO Yuanfei, WANG Xing. Screening and verification of genes related to ischemic stroke[J]. *Journal of Clinical and Pathological Research*, 2020, 40(9): 2325-2331. doi: 10.3978/j.issn.2095-6959.2020.09.017