

论 著

鉴定胶质母细胞瘤不同影像学表型核心调控基因的临床研究*

李彦国¹ 庞舒桐^{1,2} 鲍欣然³
张俊¹ 尹雪梅¹ 李文菲^{1,*}

1.秦皇岛市第一医院影像科

(河北 秦皇岛 066000)

2.河北北方学院研究生院

(河北 张家口 075000)

3.秦皇岛市第一医院神经内科

(河北 秦皇岛 066000)

【摘要】目的 本研究旨在验证MRI影像组学模型用于预测异柠檬酸脱氢酶(isocitrate dehydrogenase, IDH)野生型胶质母细胞瘤(glioblastoma, GBM)患者的预后,拟探索胶质瘤影像学异质性的基因调控机制。方法 提取130名GBM患者瘤内、瘤周全区域MRI影像组学特征,基于影像组学评分将患者分为高风险组与低风险组,利用Kaplan-Meier曲线比较不同影像组学风险组的总体生存率(overall survival, OS)。从癌症图谱(the cancer genome atlas, TCGA)数据库进行加权基因共表达网络分析(weighted gene co-expression network analysis, WGCNA),对核心基因进行功能富集分析,构建蛋白质-蛋白质相互作用(protein-protein interaction, PPI)网络,筛选出核心基因。结果 影像组学评分高风险组患者预后较差。WGCNA显示turquoise模块与高风险组相关性最强。通路分析结果提示重叠基因主要涉及表皮发育、包含胶原的细胞外基质、受体配体激活、信号受体激活剂活性及正性调控上皮细胞形态等通路。PPI网络筛选出10个与影像学异质性相关的核心基因。结论 本影像基因组学研究验证MRI的影像组学模型可用于预测IDH野生型GBM患者的预后,并提示10个核心基因可能对GBM影像学异质性具有调控作用,有望为患者治疗提供潜在分子靶点。

【关键词】胶质母细胞瘤;影像组学;异质性;加权基因共表达网络分析;基因调控

【中图分类号】R445

【文献标识码】A

【基金项目】秦皇岛市科技计划项目(202301A279);河北省医学科学研究计划(20250236)

DOI:10.3969/j.issn.1672-5131.2025.08.006

Clinical Study on Identification of Hub Gene Associated with the Differential Radiological Subtype of Glioblastoma*

LI Yan-guo¹, PANG Shu-tong^{1,2}, BAO Xin-ran³, ZHANG Jun¹, YIN Xue-mei¹, LI Wen-fei^{1,*}.

1.Department of Medical Imaging, the First Hospital of Qinhuangdao, Qinhuangdao 066000, Hebei Province, China

2.Graduate School, Hebei North University, Zhangjiakou 075000, Hebei Province, China

3.Department of Neurology, the First Hospital of Qinhuangdao, Qinhuangdao 066000, Hebei Province, China

ABSTRACT

Objective To establish a magnetic resonance imaging (MRI) radiomic model for predicting the prognosis of wild-type isocitrate dehydrogenase (IDH) glioblastoma (GBM) patients, as well as to reveal the gene regulatory mechanisms underlying glioma radiomic heterogeneity. **Methods** Radiomic features were extracted from MRI images of 130 GBM patients. Patients were divided into high-risk and low-risk groups determined by radiomic score. Kaplan-Meier curves were utilized to compare the overall survival (OS) of patients in different radiomic risk groups. The weighted gene co-expression network analysis(WGCNA) was carried out from the cancer genome atlas(TCGA) database, the functional enrichment analysis of the hub genes was carried out, the protein-protein interaction (PPI) networks of the overlapping genes were constructed to screen hub genes. **Results** The high-risk group with radiomics scores had significantly poorer survival. The turquoise module showed the strongest association with the risk groups in WGCNA. Pathway analyses revealed that the overlapping genes were primarily enriched in epidermal development, extracellular matrix containing collagen, receptor ligand activity, signal transducer activity, and positive regulation of epithelial cell morphogenesis. The PPI network screened 10 hub genes associated with imaging heterogeneity. **Conclusion** This radiogenomic study verified that MRI radiomic model can be used to predict the prognosis of wild-type IDH GBM patients, and identified that 10 hub genes may regulate imaging heterogeneity. The above genes could be used as potential therapeutic targets of GBS.

Keywords: Glioblastoma; Imaging Omics; Heterogeneity; WGCNA; Gene Regulation

异柠檬酸脱氢酶(isocitrate dehydrogenase, IDH)野生型胶质母细胞瘤(glioblastoma, GBM)主要表现为肿瘤快速浸润性生长;尽管进行了积极的多模式治疗,但5年生存率仍小于10%^[1]。组织学与分子异质性导致GBM患者预后差,并且缺乏有效的靶向治疗药物^[2]。既往多项研究证实MRI影像组学是可替代病理基因组学的重要技术,可以促进以非侵入性与可重复方式对肿瘤及其微环境的不同方面进行体内表征^[3],基于MRI影像组学特征构建的影像组学评分是GBM患者预后的独立危险因素,高风险组预后明显较差^[4-5]。然而既往研究并未关注影像学异质性的基因调控机制,一定程度上严重限制了MRI影像组学在临床工作中的推广与应用。

加权基因共表达网络分析(weighted gene co-expression network analysis, WGCNA)是把基因与表型的关联转换为基因集与临床表型的关联,将高度相关的基因归于同一表达模块,并引入相关临床特征来筛选核心基因^[6-7],为研究影像学异质性相关的核心调控基因提供了一种简便的方法^[8]。

本研究旨在验证MRI影像组学模型预测IDH野生型GBM患者预后,并探讨GBM患者影像学异质性的基因调控机制,拟寻找新的潜在治疗靶点。

1 资料与方法

1.1 数据来源 通过癌症基因组图谱数据库(The Cancer Genome Atlas, TCGA)获取197例IDH野生型GBM患者临床资料, MRI影像数据从癌症影像档案(the cancer imaging archive, TCIA)数据库下载。

纳入标准如下:(1)手术或活检后组织病理学诊断为IDH野生型胶质母细胞瘤患者;(2)术前MRI检查包括T1加权成像(T1-weighted imaging, T1WI)增强序列、T2加权成像(T2-weighted imaging, T2WI)与T2-液体反转恢复(fluid attenuated inversion recovery, FLAIR)序列。本研究临床及影像数据均来自公共数据库,符合2013年修订的《赫尔辛基宣言》要求。

1.2 影像组学评分模型构建 本研究使用软件 3D Slicer 4.11.0 从以下三部分区域提取GBM组学特征:肿瘤坏死区、肿瘤增强区及瘤周水肿区。T1WI增强图像主要识别GBM的坏死核心,通常位于肿瘤的中心区域。T2WI与T2 FLAIR序列用于识别瘤周水肿和瘤体。图像分割由副高级及以上放射科医师(具有10年神经放射经验)绘制,所有感兴趣区均使用手动勾画。基于MRI全域影像组学特征构建影像组学评分,构建评分算法如本团队前期研究所述^[4], 风险评分=特征1×系数1+特征2×系数2。

【第一作者】李彦国,男,主任医师,主要研究方向:神经影像诊断。E-mail: doctor_yan@139.com

【通讯作者】李文菲,男,主治医师,主要研究方向:神经影像诊断。E-mail: xjtulwfvip@126.com

1.3 WGCNA确定影像组学核心基因共表达模块 使用R语言的“WGCNA”包构建无标度基因共表达网络^[9]。通过Pearson相关性的分层聚类对所有基因进行聚类,选择合适的软阈值并计算无标度拟合指数与平均连接性,具有相似表达模式的基因被分为同一个共表达模块,结果与DEGs进行相关性分析,得到重叠基因。

1.4 重叠基因的功能富集分析 为探究重叠基因的生物学功能,利用R语言的“clusterprofiler”包进行GO富集分析与KEGG通路分析^[10-11],预测潜在功能及信号通路。

1.5 蛋白质-蛋白质相互作用(protein-protein interaction, PPI)网络的构建与核心基因的识别 使用STRING数据库(<http://www.string-db.org/>)构建重叠基因所编码蛋白质的PPI网络^[12],最低交互作用评分 ≥ 0.15 。结果通过Cytoscape(<http://www.cytoscape.org/>)可视化^[13]。

1.6 统计分析 分类变量采用卡方检验,使用Student's t检验或Kruskal-Wallis检验计算不同风险组连续变量的统计差异。Kaplan-Meier生存分析采用log-rank检验。双尾检验以 $P < 0.05$ 为差异具有统计学意义。

2 结果

2.1 临床特征 本研究共纳入130名GBM患者,由2名神经组放射医师在3D-slicer软件(4.8.1版本)对GBM的病灶进行识别并进行肿瘤分割(图1A),分别提取GBM瘤内、瘤周区域影像组学特征,应用10倍交叉验证的最小绝对收缩和选择算子(least absolute shrinkage and selection operator, LASSO)比例风险回归建立影像组学特征评分(图1B-C),最终8个组学特征构建影像组学评分。来根据影像组学评分阈值(6.67)将患者分为高、低风险组,其中低风险组65名患者,高风险组65名患者。不同风险组的性别、年龄、O-6-甲基鸟嘌呤-DNA-甲基转移酶(O6-methylguanine-DNA-methyltransferase, MGMT)状态与治疗方式均没有统计学差异(见表1)。结果发现不同风险组总体生存率(overall survival,

OS)存在显著差异(图1D),高风险组预后明显较差($P < 0.001$)。

2.2 WGCNA识别与高低风险组相关性最强的基因表达模块 使用R软件“limma”包分别对高、低风险组原始mRNA表达数据进行处理,获取192个DEGs。

通过R软件的“WGCNA”包分析阈值权重为1到20的网络拓扑软阈值功率,当 $\beta = 7$ 时保持了较大的基因平均连接性,以保证无标度网络的构建(图2A-B)。模块特征基因的分层聚类树状图(图2C)显示,具有相似表达模式的基因聚集成13个模块,其中turquoise模块与低风险($ME = 0.76, P = 1e-33$)呈正相关最强,与高风险($ME = -0.76, P = 1e-33$)呈负相关最强(图2D)。因此,选择turquoise模块作为关键基因模块,共9823个基因。

表1 GBM患者基线临床特征

特征	低风险组	高风险组	P
样本量	65	65	
年龄[中位数(IQR)]	58.2 (50, 63.9)	60.3 (52.4, 71.9)	
性别[n(%)]			0.585
女	25 (19.2%)	22 (16.9%)	
男	40 (30.8%)	43 (33.1%)	
MGMT[n(%)]			0.261
甲基化	21 (16.2%)	17 (13.1%)	
非甲基化	12 (9.2%)	20 (15.4%)	
未知	32 (24.6%)	28 (21.5%)	
治疗[n(%)]			0.115
放疗	12 (9.4%)	20 (15.7%)	
替莫唑胺+放疗	51 (40.2%)	44 (34.6%)	

注: MGMT, O-6-甲基鸟嘌呤-DNA-甲基转移酶, IQR, 四分位数间距。

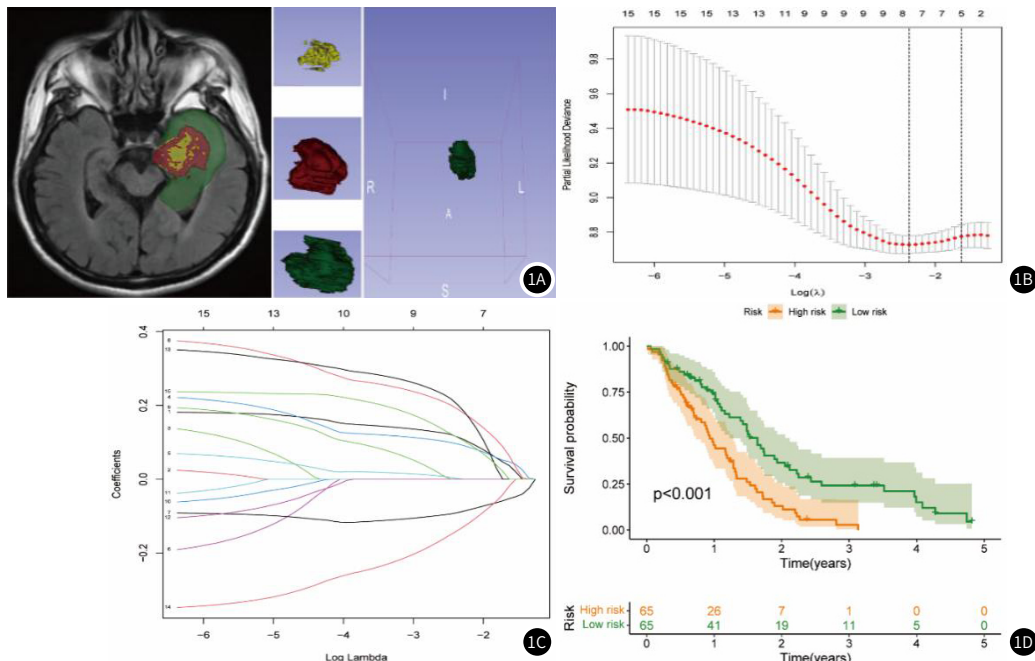


图1A~图1D 影像组学预测GBM患者预后流程图。(1A) MRI影像组学GBM病灶分割示意图,黄色代表坏死区,红色代表实质增强区,绿色代表瘤周水肿区;(1B~1C) LASSO回归降维筛选非零系数影像组学特征;(1D) Kaplan-Meier生存曲线显示高风险组预后明显差于低风险组。

2.3 重叠基因的功能分析结果 韦恩图显示turquoise基因模块与DEGs的重叠基因有67个(图3A),对重叠基因进行GO富集分析与KEGG通路分析。根据GO富集分析,生物过程(biological process, BP)模块重叠基因主要涉及表皮发育调控;细胞成分(cellular component, CC)模块重叠基因主要作用在含有胶原的细胞外基质、内质网腔、细胞-细胞连接调控;分子功能(molecular

function, MF)模块重叠基因主要富集在受体配体激活、信号受体激活剂活性的调控(图3B)。KEGG通路分析显示,重叠基因主要参与肺形态学变化、正性调控上皮细胞形态改变、调控伤害反应、正性调控肾脏发育、正性调控神经胶质细胞分化、细胞-细胞连接组织、正性调控上皮细胞增殖及表皮发育通路(图3C)。

2.4 PPI网络构建与核心基因识别 采用PPI网络分析对重叠基

因编码的蛋白质之间的相互作用进行预测与分析(图4)，整个网络包括66个节点和162条边，利用MCC算法筛选评分居前10位的

基因作为核心基因，包括SMAD3、CD34、ID1、CDH5、TEK、MEIS1、PITX2、TEAD4、VCL、EGFL7(表2)。

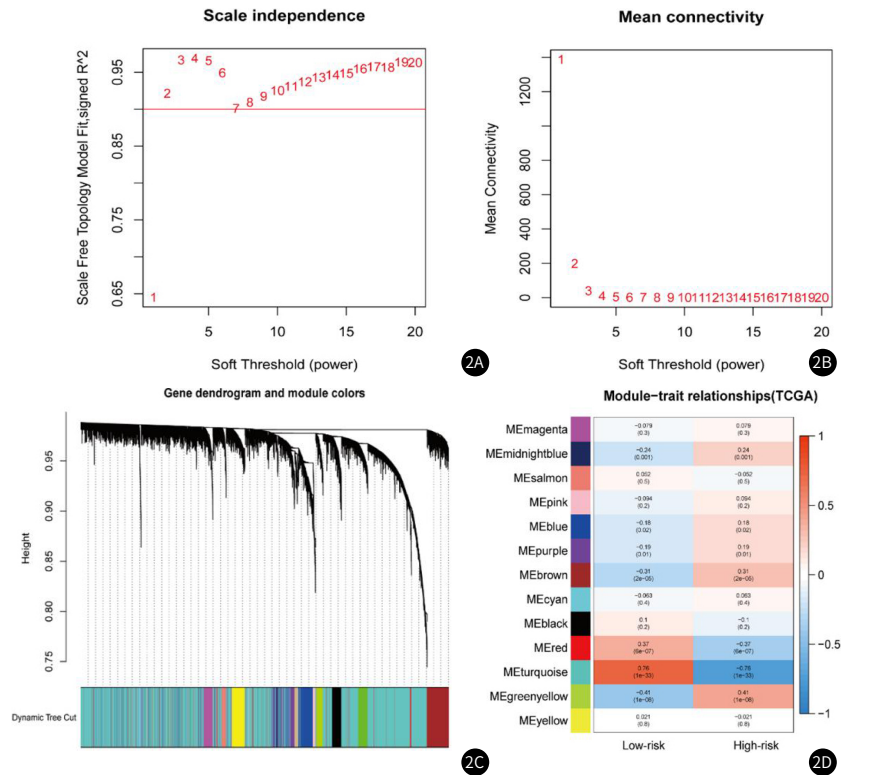


图2A ~ 图2D WGCNA分析流程。(2A)不同软阈值对应的无标度指数；(2B)不同软阈值对应的基因平均连接性；(2C)通过层次聚类构建基因树状图，树状图下方的每个色块代表一个基因模块；(2D)共表达模块特征基因与不同风险组之间的关系，颜色越深，提示相关性越高。

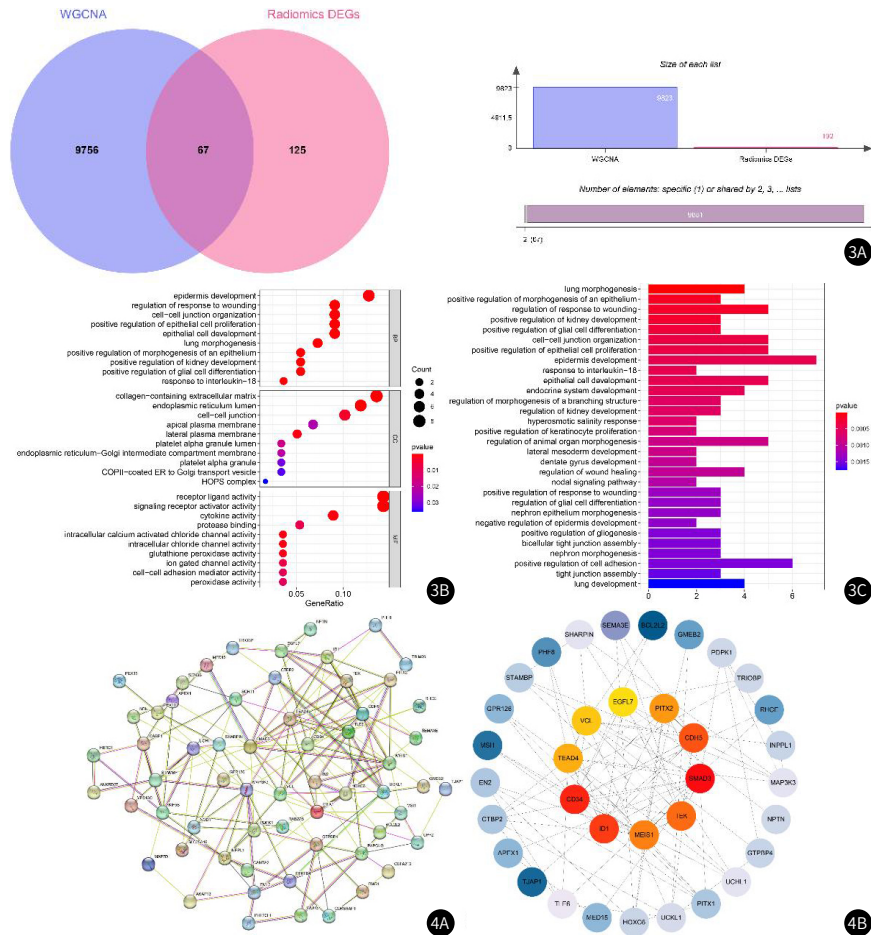


图3A ~ 图3C 差异表达基因的功能富集分析。(3A)重叠基因；(3B)重叠基因的GO富集分析；(3C)重叠基因的KEGG通路分析。图4A ~ 图4B (4A)67个重叠基因所构建的PPI网络；(4B)使用MCC算法从PPI网络中鉴定10个核心基因。

表2 通过MCC方法对网络中排名前10位的基因进行排序

等级	名称	得分
1	SMAD3	162
2	CD34	151
3	ID1	146
4	CDH5	102
5	TEK	98
6	MEIS1	97
7	PITX2	74
8	TEAD4	43
9	VCL	34
10	EGFL7	26

3 讨 论

本研究证实了影像学表征的GBM表型预后存在显著差异，高风险组患者预后不良。通过WGCNA确定了与高风险组患者相关的关键基因模块，利用MCC算法确定10个核心调控基因可能对GMB影像学异质性具有调控作用。从生物学机制上解释了影像学表型的异质性。

本研究证实基于瘤内、瘤周MRI影像学评分是IDH野生型GBM患者独立危险因素，高风险组患者预后较差，与Cepeda的研究基本一致^[14-15]。Choi^[16]回顾性收集与整理80名病理证实为胶质母细胞瘤的患者MRI影像数据，并基于5个特征构建影像学评分，有助于改善患者预后评估准确性。一项研究则通过WGCNA分析获得与GBM患者预后不良密切相关的10个相关基因，GO富集分析显示这些基因的生物学过程主要富集于细胞周期、孕酮介导的卵母细胞成熟与卵母细胞减数分裂^[17]。然而，这些研究并未将MRI的影像学模型与基因组学进行配对、揭示潜在生物学机制。

本研究筛选出的4个核心基因SMAD3、CD34、ID1、CDH5，与已发表的多项研究结果相一致。既往研究表明SMAD3在多种类型的恶性肿瘤中经常呈现低表达，具有重要的肿瘤抑制功能，SMAD3下调可增强GBM的增殖与侵袭^[18-19]。Liu等^[20]的研究也认为，SMAD3表达水平在高级别胶质瘤组织中显著降低，研究表明肿瘤血管形成过程中，CD34在血管内皮细胞与内皮祖细胞上表达增加^[21]。CD34的表达在胶质母细胞瘤不同生存期患者之间具有明显差异，且与贝伐珠单抗疗效的改善有关^[22]。DNA结合抑制剂1(ID1)基因的过表达与多种癌症的发病机制有关^[23]。已有研究发现ID1的表达与肿瘤细胞生长与存活、侵袭与迁移以及血管生成密切相关^[24]。在宫颈癌、乳腺癌以及前列腺癌中，高水平的ID1蛋白与肿瘤的低分化、恶性潜力高，以及患者的临床结局差相关^[25]。而在GBM中，抑制ID1表达可增加化疗敏感性并延缓肿瘤复发，延长患者生存期^[26]，因此针对ID1的靶向药物可能是治疗胶质母细胞瘤的一种新颖且有效的策略。

本研究方法新颖、具有一定的创新性，但仍存在诸多局限。首先，本研究基于单中心公共数据库构建影像学评分，其次，通过WGCNA获得核心调控基因，核心基因的表达还需进行进一步蛋白基础及体外试验验证；再次，不同研究表征GBM异质性影像学特征差异较大，后续可以对影响组学特征进行标准化筛选，实现影像基因组学的多维度全面研究，亟待后续的研究过程中去不断的完善与补充。

综上所述，本研究基于影像基因组学验证MRI影像学模型可以有效对GBM患者进行风险分层，从宏观到微观尺度、多角度、广维度阐明了10个核心基因可能对GMB影像学异质性具有调控作用，进一步揭示了不同影像学GBM表型的相关生物学通路，同时为探索GBM的发生、发展机制提供潜在的分子靶点。

参考文献

- [1] Ostrom Q T, Patil N, Cioffi G, et al. CBTRUS statistical report: primary brain and other central nervous system tumors diagnosed in the United States in 2013-2017[J]. *Neuro-oncology*, 2020, 22 (12 Suppl 2): iv1-iv96.
- [2] Kersch C N, Muldoon L L, Claunich C J, et al. Multiparametric magnetic resonance imaging discerns glioblastoma immune microenvironmental heterogeneity[J]. *The neuroradiology journal*, 2023: 19714009231163560.
- [3] Beig N, Bera K, Prasanna P, et al. Radiogenomic-based survival risk stratification of tumor habitat on Gd-T1w MRI is associated with biological processes in glioblastoma[J]. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, 2020, 26 (8): 1866-1876.
- [4] 李京龙, 王占秋, 顾长聪, 等. MRI放射组学特征预测胶质母细胞瘤患者预后[J]. *肿瘤影像学*. 2022, 31 (6): 596-601.
- [5] Xu P F, Li C, Chen Y S, et al. Radiomics-based survival risk stratification of glioblastoma is associated with different genome alteration[J]. *Computers in biology and medicine*, 2023, 159: 106878.
- [6] Dai Y, Sun X, Wang C, et al. Gene co-expression network analysis reveals key pathways and hub genes in Chinese cabbage (*Brassica rapa* L.) during vernalization[J]. *BMC genomics*, 2021, 22 (1): 236.
- [7] 王琦, 黄程辉, 胡一, 等. 糖酵解参与的黑色素瘤细胞放射敏感性lncRNAs的初步筛选及相关性分析[J]. *中南大学学报(医学版)* 2021, 46 (6): 565-574.
- [8] 王耀群, 陈博, 黄甫春, 等. 加权基因共表达网络分析技术在人类肿瘤研究中的应用[J]. 2019.
- [9] Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis[J]. *BMC Bioinformatics*, 2008, 9: 559.
- [10] The Gene Ontology Consortium. Expansion of the gene ontology knowledgebase and resources[J]. *Nucleic Acids Res*, 2017, 45 (D1): D331-D338.
- [11] Kanehisa M, Goto S. KEGG: Kyoto encyclopedia of genes and genomes[J]. *Nucleic Acids Res*, 2000, 28 (1): 27-30.
- [12] Szklarczyk D, Kirsch R, Koutrouli M, et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest[J]. *Nucleic Acids Res*, 2023, 51 (D1): D638-D646.
- [13] Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks[J]. *Genome Res*, 2003, 13 (11): 2498-2504.
- [14] MCGARRY S D, HURRELL S L, KACZMAROWSKI A L, et al. Magnetic resonance imaging-based radiomic profiles predict patient prognosis in newly diagnosed glioblastoma before therapy[J]. *Tomography (Ann Arbor, Mich)*, 2016, 2 (3): 223-228.
- [15] Cepeda S, Pérez-Nuñez A, García-García S, et al. Predicting short-term survival after gross total or near total resection in glioblastomas by machine learning-based radiomic analysis of preoperative MRI[J]. *Cancers*, 2021, 13 (20).
- [16] Choi Y, Ahn K J, Nam Y, et al. Analysis of heterogeneity of peritumoral T2 hyperintensity in patients with pretreatment glioblastoma: prognostic value of MRI-based radiomics[J]. *European Journal of Radiology*, 2019, 120: 108642.
- [17] Zhou J, Guo H, Liu L, et al. Construction of co-expression modules related to survival by WGCNA and identification of potential prognostic biomarkers in glioblastoma[J]. *Journal of Cellular and Molecular Medicine*, 2021, 25 (3): 1633-1644.
- [18] Jeon H Y, Pornour M, Ryu H, et al. SMAD3 promotes expression and activity of the androgen receptor in prostate cancer[J]. *Nucleic Acids Res*, 2023, 51 (6): 2655-2670.
- [19] Wu Z B, Cai L, Lin S J, et al. The miR-92b functions as a potential oncogene by targeting on Smad3 in glioblastomas[J]. *Brain Research*, 2013, 1529: 16-25.
- [20] Liu H, Yong Y, Li X, et al. Chaperone-mediated autophagy regulates cell growth by targeting SMAD3 in glioma[J]. *Neurosci Bull*, 2022, 38 (6): 637-651.
- [21] Kapoor S, Shenoy S P, Bose B. CD34 cells in somatic, regenerative and cancer stem cells: developmental biology, cell therapy, and omics big data perspective[J]. *J Cell Biochem*, 2020, 121 (5-6): 3058-3069.
- [22] Michaelsen S R, Urup T, Olsen L R, et al. Molecular profiling of short-term and long-term surviving patients identifies CD34 mRNA level as prognostic for glioblastoma survival[J]. *J Neurooncol*, 2018, 137 (3): 533-542.
- [23] Tzeng S F, De Vellis J. Id1, Id2, and Id3 gene expression in neural cells during development[J]. *Glia*, 1998, 24 (4): 372-381.
- [24] Zhao Z, Bo Z, Gong W, et al. Inhibitor of differentiation 1 (Id1) in cancer and cancer therapy[J]. *International Journal of Medical Sciences*, 2020, 17 (8): 995-1005.
- [25] Schindl M, Oberhuber G, Obermair A, et al. Overexpression of Id-1 protein is a marker for unfavorable prognosis in early-stage cervical cancer[J]. *Cancer Res*, 2001, 61 (15): 5703-5706.
- [26] Sachdeva R, Wu M, Smiljanic S, et al. ID1 is critical for tumorigenesis and regulates chemoresistance in glioblastoma[J]. *Cancer Res*, 2019, 79 (16): 4057-4071.

(收稿日期: 2024-06-07)

(校对编辑: 韩敏求、翁佳鸿)