

胶质瘤自噬相关长链非编码 RNA 的生信分析

李占成 石 铁 周卫忠

【摘要】目的 探讨胶质瘤自噬相关长链非编码 RNA(lncRNA)。方法 应用生信分析方法,首先检索 CGGA 和自噬基因列表来自人类自噬数据库(HADb),构建自噬相关 lncRNA 共表达网络;第二,多因素 Cox 回归分析胶质瘤预后相关的自噬相关 lncRNA,并根据预后 lncRNA 计算风险评分,将胶质瘤分为低危组和高危组,采用生存曲线分析风险评分与胶质瘤预后的关系;最后,应用 TCGA 数据库进一步验证。结果 CGGA 数据库筛选出 879 个 lncRNA, HADb 数据库筛选出 216 个自噬基因,构建出含 402 个 lncRNA 自噬相关 lncRNA 共表达网络;多因素 Cox 回归分析筛选出 10 个与胶质瘤病人预后有关的自噬相关 lncRNA,而且根据这些 lncRNA 计算出的风险评分是胶质瘤病人预后的独立影响因素( $P<0.05$ );生存曲线分析显示低风险组中位生存时间(1 211 d)较高风险组(346 d)明显延长( $P<0.05$ )。TCGA 数据集共纳入 160 例胶质母细胞瘤(GBM),多因素 Cox 回归分析显示,风险评分是 GBM 独立预后影响因素( $P<0.05$ );生存曲线分析显示低风险组中位生存时间(468 d)较高风险组(385 d)明显延长( $P<0.05$ )。结论 多个自噬相关 lncRNA 与胶质瘤预后相关,提示调控自噬相关 lncRNA 对胶质瘤具有潜在的治疗作用。

【关键词】胶质瘤;细胞自噬;长链非编码 RNA;预后

【文章编号】1009-153X(2021)11-0839-03 【文献标志码】A 【中国图书资料分类号】R 739.41; Q 786

Autophagy-related long non-coding RNA signature in glioma patients based on bioinformatics analysis

LI Zhan-cheng, SHI Tie, ZHOU Wei-zhong. Department of Neurosurgery, Chengde Hospital of Traditional Chinese Medicine, Chengde, 067000, China

【Abstract】Objective To explore the autophagy-related long non-coding RNA (lncRNA) signature in the glioma patients. Methods Using the biosynthesis methods: firstly, the CGGA and HADb databases were searched to construct the autophagy-related lncRNA co-expression network; secondly, multivariate Cox regression analysis was used to analyze the relationship of autophagy-related lncRNA and the prognosis of glioma patients, and then the risk score was calculated according to the lncRNA and the glioma patients were divided into low-risk group and high-risk group; thirdly, the survival curve was used to analyze the relationship between the risk score and the prognosis of glioma patients; finally, the TCGA database was used for further verification. Results Eight hundred and seventy-nine lncRNAs were screened out from the CGGA database and 216 autophagy genes were screened out from the HADb database. A co-expression network including 402 lncRNA autophagy-related lncRNAs was constructed. Multivariate Cox regression analysis showed 10 lncRNAs related to the prognosis of glioma patients. The risk score calculated based on these 10 autophagy-related lncRNAs was an independent factor influencing the prognosis of glioma patients ( $P<0.05$ ). Survival curve analysis showed that the median survival time (1 211 d) of the low-risk group was significantly longer than that (346 d) of the higher-risk group ( $P<0.05$ ). A total of 160 glioblastoma (GBM) patients were searched out from the TCGA database; multivariate Cox regression analysis showed that risk score was an independent prognostic factor of GBM patients ( $P<0.05$ ); survival curve analysis showed that the median survival time of the low-risk group (468 d) was significantly longer than that (385 d) of the higher-risk group ( $P<0.05$ ). Conclusions Many autophagy-related lncRNAs are related to the prognosis of glioma patients, suggesting that regulating autophagy-related lncRNAs has a potential therapeutic effect on glioma patients.

【Key words】Glioma; Autophagy; Long non-coding RNA; Prognosis

胶质瘤是最常见的恶性原发性中枢神经系统肿瘤,预后较差<sup>[1]</sup>,其中胶质母细胞瘤(glioblastoma multiforme,GBM)预后最差,中位总生存期在15~23个月,5年生存率低于6%<sup>[2]</sup>。

研究表明,细胞自噬在神经退行性疾病、衰老和肿瘤发病过程中具有重要作用<sup>[3]</sup>。近年来,有学者试图通过调控细胞自噬寻找胶质瘤新的潜在治疗靶点<sup>[4]</sup>。研究发现,长链非编码 RNA(long non-coding RNA, lncRNA)具有广泛的生物学功能<sup>[5]</sup>,例如调控细胞自噬相关蛋白的表达<sup>[6]</sup>。本文利用生物信息学方法分析胶质瘤细胞自噬相关 lncRNA 的表达网络,为胶质瘤的诊治提供参考。

doi:10.13798/j.issn.1009-153X.2021.11.007  
作者单位:067000 河北承德,承德医学院附属医院(承德市中医院)神经外科(李占成、石 铁、周卫忠)

1 资料和方法

1.1 胶质瘤数据集的获取 基因组图谱(CGGA, <http://www.cgga.org.cn/>)微阵列被用作训练集,以建立胶质瘤的自噬相关 lncRNA 特征,共纳入 325 例胶质瘤病人 CGGA mRNA 以及相关的临床数据。从癌症基因组图谱(TCGA, <https://cancergenome.nih.gov/>)网站下载病人的临床信息,用于验证风险评分。TCGA 数据集纳入 160 例 GBM 病人的验证数据集。

1.2 lncRNA 和自噬基因的筛选 从 CGGA mRNAseq 数据集获得 lncRNA 和自噬基因的谱。自噬基因列表来自人类自噬数据库(HADb, <http://autophagy.lu/clustering/index.html>)。所有 mRNA 表达数据通过 log2 转换进行标准化。应用 Pearson 相关性计算 lncRNA 与自噬相关基因的相关性。相关系数 $|R2|>0.3$ 的 lncRNA 被认为是自噬相关 lncRNA。

1.3 胶质瘤病人预后影响因素分析 采用单变量和多变量 Cox 回归分析评估自噬相关 lncRNA 与胶质瘤病人预后的关系。将单变量分析  $P<0.01$  的 lncRNA 纳入多变量 Cox 回归分析,根据结果以确定风险评分=基因 1×基因 1 的表达量+基因 2×基因 2 的表达量+ ... +基因 N×基因 N 的表达量。将胶质瘤病人根据中位风险评分分为高风险组和低风险组。

2 结果

2.1 自噬相关 lncRNA 共表达网络的构建 CGGA 数据集共鉴定 878 个 lncRNA, HADb 数据库共筛选出 215 个自噬相关基因。构建一个自噬相关 lncRNA 共表达网络,识别自噬相关 lncRNA,最后,筛选出 402 个 lncRNA。

2.2 胶质瘤自噬相关 lncRNA 的特征 根据自噬相关 lncRNA 共表达网络中 402 个自噬相关 lncRNA,利用 CGGA 数据集 325 例胶质瘤病人, Cox 回归分析显示 10 个自噬相关 lncRNA 是胶质瘤病人预后的独立影响因素,其中 5 个 lncRNA 是危险因素(TP53TG1、ZNF674- AS1、COX10- AS1、DDX11- AS1、SBF2- AS1), 5 个是预后保护因素(PCBP1- AS1、DHRS4- AS1、 GABPB1- AS1、 MAPKAPK5- AS1、 MIR4453HG)。见表 1、2。

2.3 自噬相关 lncRNA 与胶质瘤预后的关系 根据中位风险评分将胶质瘤病人分为低风险组和高风险组。多因素 Cox 回归分析显示,风险评分是胶质瘤预后独立影响因素( $P<0.05$ ,表 3)。生存曲线分析显示低风险组中位生存时间(1 211 d)较高风险组

(346 d)明显延长( $P<0.05$ ;图 1)。

2.4 TCGA 数据集验证结果 共纳入 160 例 GBM,根据风险评分中值分为高危组和低危组。多因素 Cox

表 1 胶质瘤自噬相关 lncRNA

lncRNA	自噬基因	相关系数	P 值
PCBP1-AS1	CCL2	-0.360	<0.001
TP53TG1	ATG2B	0.305	<0.001
DHRS4-AS1	KIF5B	0.306	<0.001
ZNF674-AS1	ATG2A	0.312	<0.001
MAPKAPK5-AS1	WDFY3	0.318	<0.001
COX10-AS1	MLST8	0.322	<0.001
GABPB1-AS1	PIK3C3	0.325	<0.001
DDX11-AS1	TSC1	0.330	<0.001
SBF2-AS1	TSC2	0.333	<0.001
MIR4453HG	NCKAP1	0.340	<0.001

表 2 胶质瘤自噬相关 lncRNA 与胶质瘤病人生存预后的多因素 Cox 回归分析

lncRNA	P 值	风险比(95%置信区间)
PCBP1-AS1	0.049	0.696(0.485~0.998)
TP53TG1	0.006	1.558(1.139~2.131)
DHRS4-AS1	0.01	0.776(0.639~0.942)
ZNF674-AS1	0.024	1.565(1.06~2.31)
MAPKAPK5-AS1	0.009	0.527(0.326~0.852)
COX10-AS1	< 0.001	2.29(1.565~3.351)
GABPB1-AS1	0.009	0.668(0.494~0.904)
DDX11-AS1	0.021	1.344(1.046~1.726)
SBF2-AS1	0.036	1.143(1.009~1.295)
MIR4453HG	< 0.001	0.577(0.426~0.781)

表 3 胶质瘤病人生存预后影响因素的多因素 Cox 回归分析

影响因素	P 值	风险比(95%置信区间)
性别	0.921	0.978(0.635~1.507)
年龄	0.987	1(0.98~1.021)
WHO 分级	< 0.001	2.296(1.678~3.143)
放疗	< 0.001	0.394(0.263~0.589)
化疗	0.023	0.626(0.417~0.938)
IDH 突变	0.01	0.528(0.324~0.858)
TP53 阳性	0.072	0.716(0.498~1.031)
EGFR 阳性	0.437	0.848(0.559~1.285)
ATRX 阳性	0.046	0.442(0.198~0.986)
EZH2 阳性	0.078	1.612(0.948~2.742)
风险评分	< 0.001	2.736(1.699~4.405)

注: IDH. 异柠檬酸脱氢酶; EGFR. 表皮生长因子受体; ATRX. a-地中海贫血/精神发育迟滞综合征 X 染色体相关基因; EZH2. zeste 基因增强子同源物 2

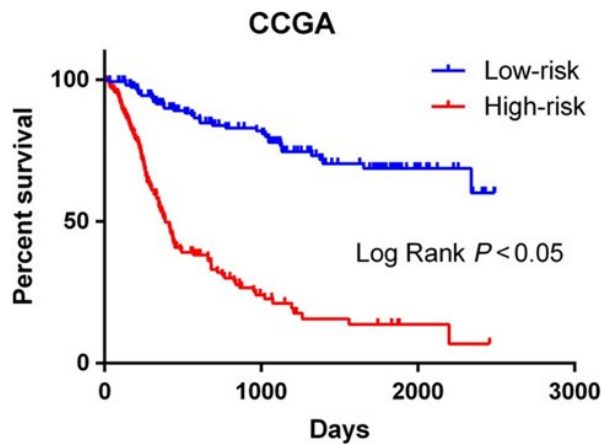


图1 生存曲线分析CGGA数据库胶质瘤病人生存期与风险评分的关系  
根据10个胶质瘤相关lncRNA计算风险评分,以中位数为准分为高风险和低风险

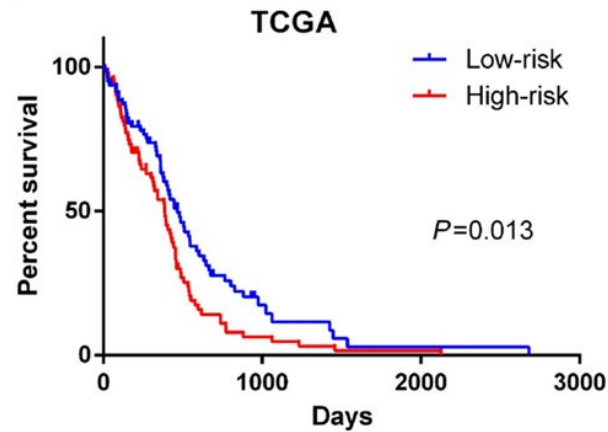


图2 生存曲线分析TCGA数据库胶质母细胞瘤病人生存期与风险评分的关系  
根据10个胶质瘤相关lncRNA计算风险评分,以中位数为准分为高风险和低风险

回归分析显示,风险评分是胶质瘤预后独立影响因素,生存曲线分析显示低风险组中位生存时间(468 d)较高风险组(385 d)明显延长( $P<0.05$ ;图2)。

3 讨论

随着胶质瘤诊疗技术的进步,很多预后影响因素得到证实,包括肿瘤切除程度、肿瘤分级等。目前,高通量测序技术被广泛地用于检测miRNA或基因的改变,可以很好地预测肿瘤复发和肿瘤转移<sup>[7]</sup>。lncRNA作为基因或miRNA的补充,为预测胶质瘤复发的风险提供了新的靶标<sup>[18]</sup>。目前,自噬在肿瘤治疗中的作用知之甚少。最近的研究表明,调节自噬将成为胶质瘤治疗的新方法<sup>[19,20]</sup>。本文利用两个数据集(CGGA和TCGA)分析自噬相关lncRNA与胶质瘤预后的关系,首先筛选出402个自噬相关lncRNA;然后,Cox回归分析显示10个自噬相关lncRNA与胶质瘤病人预后有关;再根据lncRNA的中位风险评分将胶质瘤病人分为高风险和低风险组,生存曲线分析发现低风险组中位生存时间较高风险组明显延长。

总之,我们应用生信分析方法筛选出10个与胶质瘤自噬相关的lncRNA,与胶质瘤病人预后密切相关。

【参考文献】

[1] Reni M, Mazza E, Zanos S, *et al.* Central nervous system gliomas [J]. Crit Rev Oncol Hematol, 2017, 113: 213–234.

[2] Shergais A, Bankhead A, 3rd, Luesakul U, *et al.* Current challenges and opportunities in treating glioblastoma [J]. Pharmacol Rev, 2018, 70(3): 412–445.

[3] Levine B, Kroemer G. Autophagy in the pathogenesis of disease [J]. Cell, 2008, 132(1): 27–42.

[4] Lefranc F, Facchini V, Kiss R. Proautophagic drugs: a novel means to combat apoptosis-resistant cancers, with a special emphasis on glioblastomas [J]. Oncologist, 2007, 12(12): 1395–1403.

[5] Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions [J]. Nat Rev Genet, 2009, 10 (3): 155–159.

[6] Van T Veer LJ, Rutgers EJT, Piccart M. 70-gene signature in early-stage breast cancer [J]. N Engl J Med, 2016, 375 (22): 2200–2201.

[7] Lerebours F, Cizeron-Clairac G, Susini A, *et al.* miRNA expression profiling of inflammatory breast cancer identifies a 5-miRNA signature predictive of breast tumor aggressiveness [J]. Int J Cancer, 2013, 133(7): 1614–1623.

[8] Niknafs YS, Han S, Ma T, *et al.* The lncRNA landscape of breast cancer reveals a role for DSCAM-AS1 in breast cancer progression [J]. Nat Commun, 2016, 7: 12791.

[9] Pavlides S, Vera I, Gandara R, *et al.* Warburg meets autophagy: cancer-associated fibroblasts accelerate tumor growth and metastasis via oxidative stress, mitophagy, and aerobic glycolysis [J]. Antioxid Redox Signal, 2012, 16(11): 1264–84.

(2021-08-24修回,2021-09-27修回)