

消化系统肿瘤

DOI: 10.13406/j.cnki.cyx.002574

GNG4、TIMP1 在结肠癌诊断及预后价值的生物信息学分析

刘 萌¹, 刘蔚东², 张 洁¹, 钟白云¹
(中南大学湘雅医院 1. 检验科; 2. 普外科, 长沙 410008)

【摘要】目的:通过生物信息学方法,筛选可能与结肠癌诊断及预后相关的基因。**方法:**从基因表达共享数据库(the expression omnibus, GEO)下载结肠癌相关基因芯片 GSE37364、GSE41328 的 cRNA 表达谱数据集并筛选差异表达基因(differentially expressed genes, DEGs),分别通过 STRING 在线数据库、Cytoscape 软件、GEPIA 在线工具及 Metascape 在线软件进行 DEGs 的蛋白互作(protein-protein-interactions, PPI)网络分析、核心基因筛选、预后分析、表达水平验证及通路分析。**结果:**鉴定出核心基因 18 个(*THY1*, *BMP4*, *CD44*, *CXCL8*, *GNG4*, *COL1A2*, *IL1B*, *SERPINE1*, *IGF1*, *SPPI*, *COL1A1*, *HGF*, *POSTN*, *TIMP1*, *CXCL12*, *GCG*, *IL6*, *PTGS2*),其中 *GNG4* 和 *TIMP1* 的高表达可能与预后不良相关,并参与 *PI3K-AKT*、*HIF-1*、趋化因子等信号通路的调节。**结论:**本研究为结肠癌诊断及预后研究提供了潜在靶基因,也为解析其分子调控机制提供了线索。

【关键词】结肠癌;生物信息学;差异表达基因;诊断;预后

【中图分类号】R735.3+5

【文献标志码】A

【收稿日期】2019-08-01

Bioinformatics analysis of GNG4 and TIMP1 on diagnostic and prognosis significance in colon cancer

Liu Meng¹, Liu Weidong², Zhang Jie¹, Zhong Baiyun¹

(1. Department of Clinical Laboratory; 2. Department of General Surgery,
Xiangya Hospital of Central South University)

【Abstract】Objective: To screen the gene that may be associated with diagnosis and prognosis of colon cancer through bioinformatics. **Methods:** The cRNA expression profiles data set of GSE37364 and GSE41328 were obtained from the expression omnibus(GEO) and the differentially expressed genes(DEGs) between colon adenocarcinoma tissues and normal colon tissues were screened. STRING online database, Cytoscape, GEPIA online tool and Metascape online software were used to perform the DEG-associated protein-protein-interactions(PPI) network analysis, to screen the hub gene, to perform the prognosis analysis and to perform the pathway analysis, respectively. **Results:** A total of 18 hub genes were identified(*THY1*, *BMP4*, *CD44*, *CXCL8*, *GNG4*, *COL1A2*, *IL1B*, *SERPINE1*, *IGF1*, *SPPI*, *COL1A1*, *HGF*, *POSTN*, *TIMP1*, *CXCL12*, *GCG*, *IL6* and *PTGS2*), among which high expression of *GNG4* and *TIMP1* was possible to have association with poor survival and participate in the regulation of *PI3K-AKT*, *HIF-1* and chemokine. **Conclusion:** Our study provides some potential target genes for the diagnosis and prognosis of colon cancer and gives a clue to molecular regulation mechanism.

【Key words】 colon cancer; bioinformatics; differentially expressed genes; diagnosis; prognosis

结直肠癌(colorectal cancer, CRC)是导致全球人类癌症相关死亡和发病的主要原因之一,现已成为全球第二大恶性肿瘤疾病^[1]。因生活习惯、饮食结构、环境因素等的变化,近年来 CRC 的发病率、死亡

率皆呈逐年上升趋势。早期 CRC 患者的 5 年生存率可高达 90%,若发生邻近器官及淋巴结的侵犯,其 5 年生存率降至 70.4%^[2]。目前针对 CRC 的临床治疗以手术病灶切除、放疗及辅助化疗为主,缺乏对 CRC 发生发展具体分子机制的足够认识,由此限制了治疗新策略的开展。因此,深入研究 CRC 的分子机制对 CRC 的诊断及治疗至关重要。基因芯片又名 DNA 微阵列,能平行、快速地检测成千上万个基因转录体的表达水平,在生物信息学分析中普遍应用

作者简介:刘 萌, Email: 741221561@qq.com,

研究方向:临床生化检验及肿瘤标志物研究。

通信作者:钟白云, Email: zbycsu@hotmail.com。

基金项目:湖南省重点研发计划资助项目(编号:2015GK3002)。

优先出版: <https://kns.cnki.net/kcms/detail/50.1046.R.20200618.1654.060.html>

(2020-06-19)

来挖掘潜在的疾病靶标。本研究通过分析 GEO 数据库中结肠癌相关的基因芯片数据,筛选癌组织与正常组织中差异表达的核心基因,并进行 KEGG 通路分析和预后分析,以期对结肠癌的治疗及预后寻找潜在的靶标。

1 材料与方法

1.1 材料

从美国国立生物信息中心 (National Center for Biotechnology Information, NCBI) 的 GEO 数据库 (www.ncbi.nlm.nih.gov/geo/) 获得结肠癌相关基因芯片 GSE37364、GSE41328 的表达谱数据集,分别包括 27 例结肠腺癌、38 例正常组织对照及 10 例结肠腺癌、10 例正常组织对照的 mRNA 表达信息。2 个芯片均为 Affymetrix HGU133 Plus 2.0,芯片数据生成于同一分析平台 GPL570[HG-U133_Plus_2]Affymetrix Human Genome U133 Plus 2.0 Array。

1.2 方法

通过 GEO2R 在线软件 (<https://www.ncbi.nlm.nih.gov/geo/geo2r/>) 分别对上述基因芯片数据进行筛选,设定阈值为调整后 $P < 0.05$ 、 $|\log FC| > 2$ (即表达水平变化倍数 > 4 倍) 确定 DEGs, 然后利用 STRING 在线数据库对所有 DEGs 构建 PPI 网络分析。将上述结果导入 Cytoscape (版本 3.6.1), 根据 Degree、Closeness、Betweenness 3 种算法获取核心基因, 并通过 GEPIA (<http://gepia.cancer-pku.cn/>) 在线分析工具对核心基因进行预后分析和表达差异验证。最后通过 Metascape (<http://metascape.org/>) 在线网站对核心基因进行 KEGG 通路分析。

2 结果

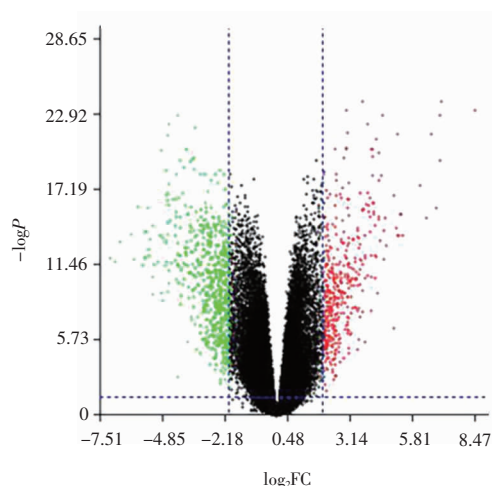
2.1 结肠腺癌与正常结肠组织差异表达基因的筛选

通过 GEO2R 的筛选, 获得 GSE37364 数据集 DEGs 775 个, 其中上调基因 311 个, 下调基因 464 个; GSE41328 数据集 DEGs 230 个, 其中上调基因 116 个, 下调基因 114 个 (图 1)。

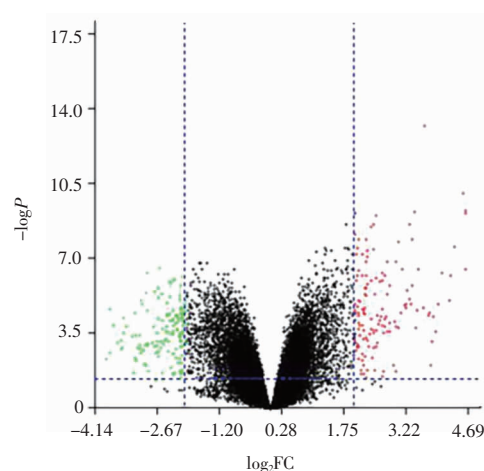
2.2 PPI 网络构建及核心基因筛选

通过 STRING 数据库对 2 个数据集所有 DEGs 进行 PPI 网络分析, 并将结果导入 Cytoscape 分析软件, 获得 671 个节点、4 042 对 PPI 关系, 169 个节点未被纳入网络分析。通过 Cytohubba 插件, 分别根据 Degree、Closeness、Betweenness 3 种算法获得前 40 位关键基因, 取交集后获得最终的关键基因共 18 个 (图 2), 分别为 *THY1*、*BMP4*、*CD44*、*CXCL8*、*GNG4*、*COLIA2*、*IL1B*、*SERPINE1*、*IGF1*、*SPPI*、*COLIA1*、*HGF*、*POSTN*、*TIMP1*、*CXCL12*、*GCG*、*IL6*、*PTGS2*。经 GEPIA 在线分析工具对上述 18 个关键基因进行全生存期的预后分析, 提示 *GNG4*、*COLIA2*、*IL1B*、*POSTN*、*TIMP1*、*PTGS2* 的表达水平

与结肠腺癌患者的预后相关。进一步验证上述基因的表达水平, 仅 *GNG4*、*TIMP1* 在结肠腺癌及正常肠黏膜中的表达水平具有统计学差异 (图 3)。



A. GSE37364 的 DEGs 火山图



B. GSE41328 的 DEGs 火山图

图 1 基因芯片表达谱差异表达基因分布图

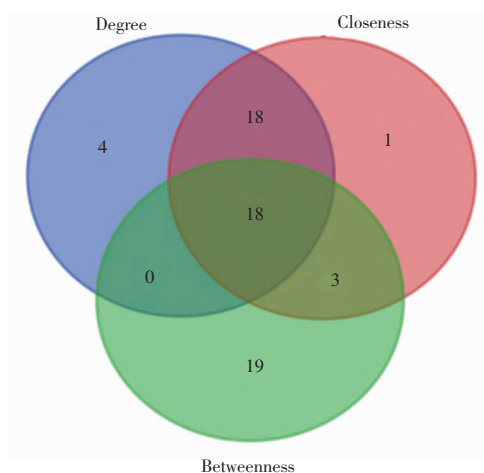
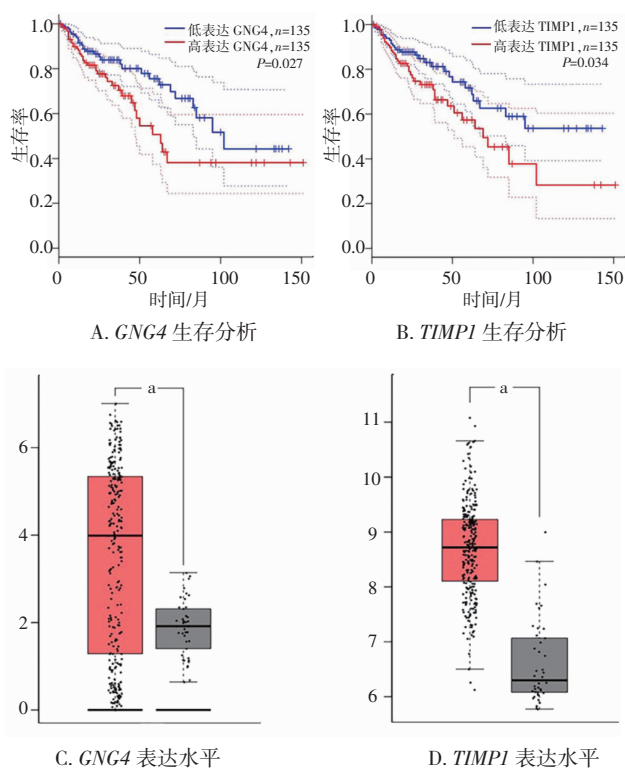


图 2 核心基因筛选韦恩图



注:红色代表癌组织($n=275$),灰色代表正常组织($n=41$);

a: $P < 0.05$, 差异有统计学意义

图3 结肠腺癌患者 *GNG4* 和 *TIMP1* 预后分析及表达水平

2.3 核心基因 KEGG 通路分析

通过 Metascape 对鉴定出的 18 个核心基因作 KEGG 通路分析,其中与 *TIMP1* 和 *GNG4* 相关的显著富集的通路见表1。

3 讨论

结肠癌的发生受多种因素影响,被认为是累积遗传、表观遗传、个体特质及内分泌异常共同作用的结果^[3]。因此,了解结肠癌发生的分子机制至关重要。生信分析作为一种挖掘疾病新靶标的有效方法,被广泛用于当前研究当中。

本研究通过分析基因芯片 GSE37364、GSE41328

的表达谱数据集,分别得到了 775 和 230 个差异表达基因。将所有差异表达基因纳入 Cytoscape 的网络节点分析筛选核心基因,共得到 18 个关键基因。经 GEPIA 数据库验证,*GNG4*、*COL1A2*、*IL1B*、*POSTN*、*TIMP1*、*PTGS2* 的表达水平与结肠癌患者的预后相关。*GNG4*、*COL1A2*、*POSTN* 和 *TIMP1* 较高的表达预示较低的总体存活率,*IL1B* 和 *PTGS2* 反之。对上述基因的表达水平进一步验证,仅 *GNG4*、*TIMP1* 在正常组织与结肠癌组织中表达差异具有统计学差异。

G 蛋白又称鸟苷酸结合蛋白,由 α 、 β 、 γ 3 条蛋白链组成,*GNG4* 是 G 蛋白三聚体的 14 个 γ 亚基蛋白之一,主要在脑部及卵巢表达,在结肠组织也有表达。*GNG4* 在胶质瘤、膀胱癌中呈现高度甲基化^[4-5],尚未有研究分析 *GNG4* 在结肠癌中的甲基化水平。*GNG4* 在肿瘤中的作用具有一定的争议,其在胶质母细胞瘤中和肾癌中具有潜在的抑癌作用^[6-7];在直肠癌中,Palma P 等^[8]认为 *GNG4* 的高 mRNA 表达水平是晚期直肠癌患者对术前放化疗应答的重要预后因素;然而,与本研究结论相似,在针对其他结肠癌基因芯片的表达谱中,*GNG4* 也作为中枢基因,其表达水平高于正常结肠组织,且高表达预示较差的预后^[9],这可能与 *GNG4* 相关的化疗药物耐药性相关^[10-11]。18 个中枢基因的 KEGG 通路分析显示,*GNG4* 与 *CXCL12*、*CXCL8* 共同富集于趋化因子信号通路。*CXCR4* 是 *CXCL12* 的特异性受体,在胶质母细胞瘤中可能是 *GNG4* 的潜在靶标^[6],而通过 *CXCR4* 可刺激 $G\alpha$ -Rho 信号轴介导细胞迁移^[12]。总体而言,*GNG4* 在结肠癌中的具体影响及作用机制尚无研究,值得进一步探讨。

细胞外基质是肿瘤相关微环境的重要组成部分之一^[13],对于组织极性和构架的保持有重要作用。重构的细胞外基质促进上皮间充质转化,从而加速肿瘤侵袭和转移的进程^[14]。*TIMP1* 即金属蛋白酶组

表1 核心基因 KEGG 通路分析

信号通路	P 值	数目	基因
Hsa05200: 肿瘤通路	1.81E-10	8	<i>IGF1</i> 、 <i>IL6</i> 、 <i>CXCL8</i> 、 <i>PTGS2</i> 、 <i>BMP4</i> 、 <i>CXCL12</i> 、 <i>HGF</i> 、 <i>GNG4</i>
Hsa04151: <i>PI3K</i> - <i>AKT</i> 信号通路	2.99E-9	7	<i>COL1A1</i> 、 <i>COL1A2</i> 、 <i>SPP1</i> 、 <i>HGF</i> 、 <i>IGF1</i> 、 <i>GNG4</i> 、 <i>IL6</i>
Hsa04371: <i>Apelin</i> 信号通路	1.40E-4	3	<i>SPP1</i> 、 <i>GNG4</i> 、 <i>SERPINE1</i>
Hsa04066: <i>HIF-1</i> 信号通路	8.41E-7	4	<i>TIMP1</i> 、 <i>IGF1</i> 、 <i>IL6</i> 、 <i>SERPINE1</i>
Hsa04062: 趋化因子信号通路	3.16E-4	3	<i>CXCL12</i> 、 <i>CXCL8</i> 、 <i>GNG4</i>
Hsa04014: <i>Ras</i> 信号通路	6.03E-4	3	<i>GNG4</i> 、 <i>IGF1</i> 、 <i>HGF</i>

组织抑制因子 1,可抑制金属蛋白酶对细胞外基质的降解,调节细胞基质重构,从而抑制肿瘤的转移。与正常组织相比,*TIMP1* 在结肠癌组织中的表达水平上升^[15]。诊断原发性结直肠癌时,*TIMP1* 与 *CEA* 具有良好的一致性,ROC 曲线分析其曲线下面积达 0.75^[16]。此外,*TIMP1* 作为一种分泌蛋白,能在血浆等体液标本中检测到,有利于临床检测项目的开展。因此,*TIMP1* 作为结直肠癌的诊断标志物具有广阔前景。研究发现,*TIMP1* 的表达与淋巴结转移、远处转移、血管侵犯和肿瘤分化程度相关,表明 *TIMP1* 的高表达与结肠癌的不良预后相关^[17-18]。*TIMP1* 除了作为金属蛋白酶抑制剂起主要作用,亦具有抑制结肠癌细胞增殖和抗凋亡的作用,可通过 FAK-PI3K/AKT 和 MAPK 信号通路影响肿瘤的进展。进一步的基因富集分析鉴定出 *TIMP1* 相关的生物学现象及通路,如上皮间充质转化、凋亡、血管生成、*KRAS*、低氧等^[17]。与本研究的结果相似,Wu Z 等^[19]也在结直肠癌中鉴定出 *TIMP1* 与 *HIF-1* 信号通路相关。*HIF-1* 是缺氧诱导因子,涉及肿瘤的代谢、增殖、凋亡、耐药等多个方面。在其他组织类型的细胞中,*TIMP1* 可作为转录因子 *HIF-1* 的靶基因,沉默 *HIF-1 α* 后其表达水平下降^[20-21]。然而在结肠癌中,*TIMP1* 与 *HIF-1* 相互作用的生物学效应和分子机制尚需更深入的研究。

综上所述,*GNG4* 和 *TIMP1* 可能是结肠癌诊断及预后良好的生物标志物,可为结肠癌的研究提供方向。然而,尚需进一步的分子生物学实验来确定其具体功能和作用机制。

参 考 文 献

- [1] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2017[J]. CA Cancer J Clin, 2017, 67(1): 7-30.
- [2] Miller KD, Siegel RL, Lin CC, et al. Cancer treatment and survivorship statistics, 2016[J]. CA Cancer J Clin, 2016, 66(4): 271-289.
- [3] Sakai E, Fukuyo M, Ohata K, et al. Genetic and epigenetic aberrations occurring in colorectal tumors associated with serrated pathway[J]. Int J Cancer, 2016, 138(7): 1634-1644.
- [4] Shukla S, Pia Patric IR, Thinagararjan S, et al. A DNA methylation prognostic signature of glioblastoma: identification of NPTX2-PTEN-NF- κ B nexus[J]. Cancer Res, 2013, 73(22): 6563-6573.
- [5] Zhang Y, Fang L, Zang Y, et al. Identification of core genes and key pathways via integrated analysis of gene expression and DNA methylation profiles in bladder cancer[J]. Med Sci Monit, 2018, 24: 3024-3033.
- [6] Pal J, Patil V, Mondal B, et al. Epigenetically silenced *GNG4* inhibits SDF1 α /CXCR4 signaling in mesenchymal glioblastoma[J]. Genes Cancer, 2016, 7(3-4): 136-147.
- [7] Maina EN, Morris MR, Zatyka M, et al. Identification of novel VHL target genes and relationship to hypoxic response pathways[J]. Oncogene, 2005, 24(28): 4549-4558.
- [8] Palma P, Cano C, Conde-Muñoz R, et al. Expression profiling of rectal tumors defines response to neoadjuvant treatment related genes[J]. PLoS One, 2014, 9(11): e112189.
- [9] Yang W, Ma J, Zhou W, et al. Identification of hub genes and outcome in colon cancer based on bioinformatics analysis[J]. Cancer Manag Res, 2018, 11: 323-338.
- [10] Yu C, Hong H, Lu J, et al. Prediction of target genes and pathways associated with cetuximab insensitivity in colorectal cancer[J]. Technol Cancer Res Treat, 2018, 17: 1-13.
- [11] 王旦旦, 吕琳, 陈智如, 等. 顺铂耐药相关基因的筛选、验证及生物信息学分析[J]. 重庆医学, 2019, 48(1): 124-128.
- [12] Tan W, Martin D, Gutkind JS. The α 13-Rho signaling axis is required for SDF-1-induced migration through CXCR4[J]. J Biol Chem, 2006, 281(51): 39542-39549.
- [13] Lu P, Weaver VM, Werb Z. The extracellular matrix: a dynamic niche in cancer progression[J]. J Cell Biol, 2012, 196(4): 395-406.
- [14] Radisky ES, Radisky DC. Matrix metalloproteinase-induced epithelial-mesenchymal transition in breast cancer[J]. J Mammary Gland Biol Neoplasia, 2010, 15(2): 201-212.
- [15] Zeng C, Chen Y. *HTR1D*, *TIMP1*, *SERPINE1*, *MMP3* and *CNR2* affect the survival of patients with colon adenocarcinoma[J]. Oncol Lett, 2019, 18(3): 2448-2454.
- [16] Nielsen HJ, Brünner N, Jorgensen LN, et al. Plasma *TIMP-1* and *CEA* in detection of primary colorectal cancer: a prospective, population based study of 4 509 high-risk individuals[J]. Scand J Gastroenterol, 2011, 46(1): 60-69.
- [17] Song G, Xu S, Zhang H, et al. *TIMP1* is a prognostic marker for the progression and metastasis of colon cancer through FAK-PI3K/AKT and MAPK pathway[J]. J Exp Clin Cancer Res, 2016, 35(1): 148.
- [18] Giaginis C, Nikiteas N, Margeli A, et al. Serum tissue inhibitor of metalloproteinase 1 and 2 (*TIMP-1* and *TIMP-2*) levels in colorectal cancer patients: associations with clinicopathological variables and patient survival[J]. Int J Biol Markers, 2009, 24(4): 245-252.
- [19] Wu Z, Liu Z, Ge W, et al. Analysis of potential genes and pathways associated with the colorectal normal mucosa-adenoma-carcinoma sequence[J]. Cancer Med, 2018, 7(6): 2555-2566.
- [20] Wang J, Ni Z, Duan Z, et al. Altered expression of hypoxia-inducible factor-1 α (*HIF-1 α*) and its regulatory genes in gastric cancer tissues[J]. PLoS One, 2014, 9(6): e99835.
- [21] Yang C, Zeng S, Lv M. Effect of shRNA inhibiting *HIF1 α* gene on *TIMP1* expression in RPE cells[J]. J Huazhong Univ Sci Technolog Med Sci, 2006, 26(1): 133-136.

(责任编辑:冉明会)