

## • 论著 •

## COL5A2 在胃癌组织中的表达及其与患者预后的关系

徐珊 张雪琳 田字彬

(青岛大学附属医院消化内科, 山东 青岛 266003)

**[摘要]** 目的 探讨 V 型胶原蛋白  $\alpha 2$ (COL5A2) 基因在胃癌(GC)组织中的表达及与患者预后的关系。方法 使用 Oncomine 数据库分析 COL5A2 基因在 GC 组织的表达;通过 cBioPortal 数据库探寻 COL5A2 基因突变情况;利用 Kaplan-Meier Plotter 平台绘制 COL5A2 mRNA 表达水平与 GC 患者预后的生存曲线;使用 GEPIA 数据库对 COL5A2 基因在 GC 中的表达及临床预后关系进行验证;采用 GeneMANIA 数据库分析 COL5A2 基因-基因互作网络,并进行功能富集分析。通过实时定量聚合酶链反应(qRT-PCR)技术检测 4 种 GC 细胞株(HGC-27、MKN-45、BGC-823 和 SGC-7901)COL5A2 mRNA 的表达水平;用小干扰 RNA(siRNA)抑制 HGC-27 细胞 COL5A2 基因的表达,将细胞分为转染 NC siRNA 的阴性对照组(si-NC 组)和转染 COL5A2 siRNA 的转染组(si-1、si-2 和 si-3 组),最后通过 CCK-8 和细胞克隆实验检测 COL5A2 基因下调对 GC 细胞增殖的影响。结果 数据库分析显示 COL5A2 基因在 GC 组织中高表达,且与 GC 患者的肿瘤分期呈正相关;cBioPortal 数据库研究发现 COL5A2 基因在 GC 中的突变率为 10.36%,并且 COL5A2 基因突变组比非突变组总生存期更长;生存曲线分析表明高 COL5A2 mRNA 表达与 GC 患者较差的临床预后显著相关;此外,通过 GeneMANIA 数据库探索发现 COL5A2 基因与 COL5A1、BMP1、ITGA1 等 20 种基因互作调控,主要与细胞外基质受体交互、骨骼系统开发以及肽交联等有关。qRT-PCR 检测结果显示,COL5A2 mRNA 在 HGC-27 和 MKN-45 细胞株中显著高表达;HGC-27 细胞经过下调 COL5A2 基因表达后,与 si-NC 组相比,si-1 和 si-2 组的细胞增殖能力以及克隆形成能力均明显下降。结论 COL5A2 基因在 GC 组织和细胞中高表达,与患者的肿瘤分期和不良预后密切相关,COL5A2 基因突变患者预后较好,且 COL5A2 基因与细胞外基质受体交互、骨骼系统开发、肽交联等有关,下调 COL5A2 基因的表达可抑制 GC 细胞的增殖和克隆形成能力,因此 COL5A2 基因有望成为治疗 GC 的潜在新靶点。

**[关键词]** 胃肿瘤;细胞系;肿瘤;胶原 V 型;数据库;遗传学;数据库挖掘;基因表达;突变;细胞增殖;预后;生物标记;肿瘤

[中图分类号] R735.2;R394

[文献标志码] A

**EXPRESSION OF COL5A2 IN GASTRIC CANCER TISSUE AND ITS PROGNOSTIC SIGNIFICANCE** XU Shan, ZHANG Xuelin, TIAN Zibin (Department of Gastroenterology, The Affiliated Hospital of Qingdao University, Qingdao 266003, China)

**[ABSTRACT]** **Objective** To investigate the expression of the collagen type V alpha 2 (COL5A2) gene in gastric cancer (GC) tissue and its prognostic significance. **Methods** The Oncomine database was used to analyze the expression of the COL5A2 gene in GC tissue. The cBioPortal platform was used for information of COL5A2 gene mutation. The Kaplan-Meier Plotter was used to draw survival curves for COL5A2 mRNA expression level and prognosis of GC patients. The GEPIA database was employed to validate the expression of COL5A2 in GC and its relationship with clinical prognosis. The COL5A2 gene-gene interaction network was constructed through the GeneMANIA database, and functional enrichment analysis was conducted in the Metascape database. qRT-PCR was performed to determine the expression levels of COL5A2 mRNA in four GC cell lines (HGC-27, MKN-45, BGC-823, and SGC-7901). Small interfering RNA (siRNA) was used to inhibit the expression of the COL5A2 gene in HGC-27 cells. The cells were divided into negative control siRNA-transfected group (si-NC group) and COL5A2 siRNA-transfected groups (si-1, si-2, and si-3 groups). The Cell Counting Kit-8 and colony formation assay were applied to detect the effect of COL5A2 gene downregulation on GC cell proliferation. **Results** Database analysis showed that the COL5A2 gene was highly expressed in GC tissue and was positively correlated with the tumor stage of patients with GC. cBioPortal database analysis found that the mutation rate of the COL5A2 gene in GC was 10.36%, and the overall survival was longer in the mutation group than in the non-mutation group. Survival curve analysis showed that high COL5A2 mRNA expression was significantly associated with a poor clinical prognosis in patients with GC. According to GeneMANIA database analysis, COL5A2 interacted with 20 genes including COL5A1, BMP1, and ITGA1, which were mainly related to extracellular matrix receptor interaction, skeletal system development, and peptide cross-linking. qRT-PCR showed that COL5A2 mRNA was highly expressed in HGC-27 and MKN-45 cell lines. After COL5A2 gene expression was downregulated in HGC-27 cells, proliferative and colony-forming capacities were significantly

decreased in the si-1 and si-2 groups compared with the si-NC group.

**Conclusion** The COL5A2 gene shows high expression in GC tissue and cells, and is closely related to tumor stage and poor prognosis of

[收稿日期] 2020-11-19; [修订日期] 2021-01-18

[基金项目] 国家自然科学基金资助项目(81802777)

[通讯作者] 田字彬, Email: tianzb@qduhospital.cn

GC patients. Patients with COL5A2 gene mutation have a better prognosis. The COL5A2 gene is associated with extracellular matrix receptor interaction, skeletal system development, and peptide cross-linking. Downregulating the COL5A2 gene can inhibit the proliferative and colony-formation abilities of GC cells. Therefore, the COL5A2 gene may be a potential new target for the treatment of GC.

[KEY WORDS] Stomach neoplasms; Cell Line, tumor; Collagen type V; Databases, genetic; Database mining; Gene expression; Mutation; Cell proliferation; Prognosis; Biomarkers, tumor

胃癌(GC)是全球五大常见恶性肿瘤之一,也是全球癌症相关死亡的三大原因之一<sup>[1]</sup>。每年中国GC新发病例数达70万,新增死亡病例数约50万,平均每天有1300多例GC患者死亡<sup>[2]</sup>。尽管目前GC的诊断和治疗均取得了显著成效,但大多数患者被发现时已是晚期或已发生转移,GC患者的5年生存率仍不足10%<sup>[3]</sup>。因此,探索GC诊断和预后的精准生物标志物十分必要。

V型胶原蛋白 $\alpha 2$ (COL5A2)属于V型胶原蛋白家族的一个亚型,是一种具有调节作用的纤维胶原,在原纤维的形成和肿瘤浸润转移的过程中起重要作用<sup>[4-6]</sup>。胶原蛋白超家族由28种不同类型的胶原组成,是细胞外基质的主要组成部分,占人体总蛋白的30%,对维持细胞正常的内稳态至关重要<sup>[7]</sup>。而细胞外基质是肿瘤微环境发生的关键,故近年来各类胶原蛋白在肿瘤中的作用受到了广泛关注。研究显示,COL5A2基因与膀胱癌、食管癌、结直肠癌以及胰腺腺癌等的发生发展具有密切的关系<sup>[8-11]</sup>。WANG<sup>[12]</sup>利用芯片数据和网络插件筛选与GC相关的差异表达基因,鉴定出COL5A2基因是GC中的一个关键基因。之后,WEI等<sup>[13]</sup>通过加权相关网络和机器学习方法发现COL5A2基因是GC的不良预后因素。但上述相关研究只停留在基因网络宏观层面,没有具体明确COL5A2基因在GC组织或细胞中的表达水平、突变情况以及可能参与的生物学过程。本研究通过对肿瘤芯片数据库的挖掘和体外GC细胞功能实验的研究,全面分析COL5A2基因在GC组织中的表达及其与患者预后的关系,探讨COL5A2基因在GC中可能发挥的功能,并评估其作为GC诊断和预后生物标志物的潜力。

## 1 材料和方法

### 1.1 利用 Oncomine 数据库分析 COL5A2 基因表达水平

登录 Oncomine 数据库(<https://www.oncomine.org/>)<sup>[14]</sup>检索COL5A2基因在不同肿瘤类型中的表达。数据设置的筛选条件如下:①基因为COL5A2;②数据比较类型为肿瘤组织与正常组织;

③临界值设定范围为 $P < 0.0001$ ,差异表达级别 $> 2$ ,基因排序为前10%。在以上条件的基础上检索COL5A2基因在GC中的表达,限定条件为:①肿瘤类型为GC;②数据类型为mRNA;③样本类型为临床标本。

### 1.2 利用 cBioPortal 数据库分析 GC 中 COL5A2 基因的突变情况

在cBioPortal数据库(<https://www.cbioportal.org/>)<sup>[15]</sup>当中,选择“Stomach Adenocarcinoma (TCGA, Firehose Legacy)”数据集,选择“带有mRNA数据(RNA Seq V2)”的病例集,输入基因名“COL5A2”,分析其在GC中的突变情况。在此基础上点击“Comparison/Survival”按钮,在Survival板块中生成COL5A2基因突变组与非突变组的生存曲线图。

### 1.3 用 Kaplan-Meier Plotter 数据库对 COL5A2 基因与 GC 患者预后的相关性进行分析

进入Kaplan-Meier Plotter平台(<https://km-plot.com/analysis/>)<sup>[16]</sup>,在肿瘤类型中选择“Gastric cancer”进行分析,生存条件限定为“overall survival”,检索获得COL5A2 mRNA与GC患者预后相关的生存曲线图。

### 1.4 利用 GEPIA 数据库验证 COL5A2 mRNA 的表达及与患者预后关系

登录GEPIA数据库(<http://gepia.cancer-pku.cn/>)<sup>[17]</sup>,点击“Expression DIY”按钮,进入“Box-plot”页面,检索COL5A2基因在GC组织中的表达水平,设置检索条件为:① $|\log_2 FC| \geq 1, P < 0.01$ ;②正常数据的匹配为Match TCGA normal and GTEx data。此外,选择“Stage plot”页面,默认原始参数,检索COL5A2基因在不同GC分期中的表达情况。而在“Survival Plots”页面中,设置组内截断值为“Median”,输入肿瘤名称“Stomach adenocarcinoma (STAD)”和基因名称“COL5A2”,获取患者生存曲线图。

### 1.5 利用 GeneMANIA 数据库绘制 COL5A2 基因互作网络图

在GeneMANIA数据库(<https://genemania.org/>)

org/)[18] 官网首页中,输入基因“COL5A2”,构建其相关基因网络树状图。之后通过 Metascape 在线网络工具(<http://metascape.org/gp/index.html>)[19],将 COL5A2 及其相关联基因输入,进行功能富集分析,限定条件为:①输入物种为 Any species;②分析物种为 H. sapiens(20)。

## 1.6 细胞培养及 GC 细胞 COL5A2 mRNA 表达水平的检测

正常人胃上皮细胞株 GES-1 以及 GC 细胞株 HGC-27、MKN-45 购自中国武汉普诺赛生命科技有限公司,GC 细胞株 BGC-823 和 SGC-7901 获赠自青岛大学病原微生物研究中心。所有细胞均培养于含体积分数 0.10 胎牛血清及 1%青链霉素混合液的 DMEM 培养基中,置于 37 °C、含体积分数 0.05 CO<sub>2</sub> 的湿化培养箱中常规培养。

采用实时定量聚合酶链反应(qRT-PCR)检测 COL5A2 mRNA 的表达水平。先利用 Trizol 试剂(美国 Invitrogen 公司)提取细胞总 RNA,以总 RNA 为模板使用反转录试剂盒(日本 TaKaRa 公司)将 RNA 逆转录为 cDNA,之后按照 TB Green Premix Ex Taq II 试剂盒(日本 TaKaRa 公司)的说明书步骤配制相应体系来进行 PCR 扩增。所涉及的引物序列见表 1。以 GAPDH 基因为内参对照,以  $2^{-\Delta\Delta CT}$  测定目的基因 mRNA 的相对表达量。

表 1 qRT-PCR 检测 COL5A2 mRNA 表达水平所用引物序列

引物名称	引物序列(5'→3')	引物长度(bp)
GAPDH	上游引物:AACAGCCTCAAGATCATCAGCAA	23
	下游引物:GACTGTGGTCATGAGTCCTTCCA	23
COL5A2	上游引物:ACGATCAAGCTAAGAACCTCAA	22
	下游引物:CACATTTCATTCCGCTTAGAG	22

## 1.7 siRNA 对 HGC-27 细胞的转染以及转染率的测定

将处于对数生长期的 HGC-27 细胞以同等密度接种于 6 孔板上,待孔内细胞融合为 60%~70% 时,使用 Lipofectamine 3000(美国 Invitrogen 公司)进行转染,使 COL5A2 基因在细胞中低表达。siRNA 由上海吉玛制药技术有限公司合成,具体序列见表 2。将细胞分为阴性对照组(si-NC 组)、转染 COL5A2-homo-4038 siRNA 组(si-1 转染组)、转染 COL5A2-homo-3838 siRNA 组(si-2 转染组)和转染 COL5A2-homo-4506 siRNA 组(si-3 转染组)。siRNA 和 Lipofectamine 3000 按 1:1 比例加入完全培养基中,培养 48 h 后通过 qRT-PCR 方法(检测

方法同 1.6) 检测各组细胞的 COL5A2 mRNA 的表达情况。

表 2 细胞转染所涉及的 siRNA 序列

siRNA 名称	siRNA 序列(5'→3')	序列长度(bp)
NC	正义链:UUCUCCGAACGUGUCACGUTT	21
	反义链:ACGUGACACGUUCGGAGAATT	21
COL5A2-homo-4038	正义链:GCACGUGUGAUGACCUAAATT	21
	反义链:UUUAGGUCAUCACACGUGCTT	21
COL5A2-homo-3838	正义链:CCACCUUACAGCUGCUCUUTT	21
	反义链:AAGAGCAGCUGUAAAGGUGGTT	21
COL5A2-homo-4506	正义链:GCUCUAAGCGGAAUGGAAATT	21
	反义链:UUUCCAUUCCGCUUAGAGCTT	21

## 1.8 CCK-8 实验检测 GC 细胞的增殖能力

将处于对数生长期的 HGC-27 细胞制成单细胞悬液,按照 2 000 个/孔细胞密度接种于 96 孔板中并进行转染。分别于转染培养后 0、24、48、72 h 于每孔加入 CCK-8 试剂(上海碧云天生物技术有限公司)10  $\mu$ L,置于恒温细胞培养箱内孵育 1.5 h,以酶标仪检测 450 nm 波长处各孔吸光度(A)值。

## 1.9 克隆形成实验检测 GC 细胞克隆形成能力

收集转染 48 h 的 HGC-27 细胞制备成单细胞悬液,以 1 000 个/孔的细胞密度接种于 6 孔板中。置于恒温培养箱中培养,每 4~5 d 更换完全培养基,待 2 周形成细胞克隆后终止培养。使用多聚甲醛溶液固定细胞 30 min 以后,用结晶紫溶液染色 15 min,用 PBS 洗涤 2~3 次,晾干后计算细胞克隆数量。

## 1.10 统计学方法

利用 SPSS 21.0 和 GraphpadPrism 7.0 软件进行统计学分析。计量资料数据以  $\bar{x} \pm s$  表示,在两个及两个以上因素影响下的多组数据比较采用析因设计的方差分析,组间两两比较采用 Turkey 检验;在单一因素影响下的多组数据均数的比较采用单因素方差分析,两组间的比较采用  $t$  检验,使用 Log-rank 检验评估患者生存差异。以  $P < 0.05$  为差异有统计学意义。

## 2 结 果

### 2.1 GC 组织样本中 COL5A2 基因的表达情况

利用 Oncomine 数据库在线分析 COL5A2 基因在多种类型肿瘤组织中的表达情况,结果显示,涉及 COL5A2 基因的研究共 417 项,78 项有统计学差异,在肿瘤组织中高表达 75 项,低表达 3 项。其中在 GC 中高表达的数据集有 7 项,包括 5 项 mRNA 和 2 项 DNA 数据集。选择 Cui Gastric、Wang

Gastric、Chen Gastric、Cho Gastric 和 DErrico Gastric 5 个 mRNA 数据集来确定 *COL5A2* 基因在 GC 中的表达情况,共涉及组织样本 478 例。将数据集进行荟萃分析显示,*COL5A2* mRNA 在 GC 差异表达基因的中位秩为 105.00,  $P$  值为  $4.05 \times 10^{-7}$ ,显示 *COL5A2* mRNA 在 GC 组织中呈高表达。以箱式图对上述 5 项研究进行分析,结果显示 *COL5A2* 基因在 GC 组织中的表达均明显高于正常胃黏膜组织。见图 1。

## 2.2 *COL5A2* 基因在 GC 组织的变异情况

在 cBioPortal 数据库中一共访问了 415 例 GC 样本信息,其中 43 例发生 *COL5A2* 基因变异(突变率为 10.36%),包括扩增 6 例,深度缺失 5 例,截断突变 5 例,错义突变 11 例,信使 RNA 水平上调 7 例,信使 RNA 水平下调 9 例。比较 *COL5A2* 基因突变样本与非突变样本的生存差异,显示 *COL5A2* 基因突变样本患者比非突变样本患者的总生存期(OS)更长。其中 *COL5A2* 基因突变组和非突变组的中位 OS 分别为 57.39、26.45 个月。见图 2A。

## 2.3 *COL5A2* 基因与 GC 患者临床预后的相关性

对 Kaplan-Meier 生存曲线分析显示,*COL5A2*

mRNA 高表达的患者的 OS 明显短于低表达的患者( $HR = 1.35, P < 0.05$ )。其中 *COL5A2* mRNA 高表达和低表达的中位 OS 分别为 26.70、44.07 个月。见图 2B。

## 2.4 *COL5A2* 基因在 GC 组织中的表达及其与患者预后的关系

在 GEPIA 数据库中对 *COL5A2* 基因在 GC 组织中的表达及其与 GC 患者预后的关系进行验证,结果显示,基于 TCGA 和 GTEx 数据集,*COL5A2* mRNA 在 408 例 GC 组织中的表达明显高于 211 例正常胃黏膜组织;同时还发现 *COL5A2* mRNA 表达水平与 GC 患者的分期有显著关系,I 期胃癌患者 *COL5A2* mRNA 表达水平比 II、III 和 IV 期 GC 患者表达低( $F = 2.65, P < 0.05$ )。此外,GEPIA 在线生存分析显示,GC 患者 *COL5A2* 基因表达水平与 OS 有关( $HR = 1.5, P < 0.05$ )。见图 2C。

## 2.5 *COL5A2* 的关联基因及其相关功能的预测

为了进一步研究 *COL5A2* 基因可能参与的生物学通路和功能,本研究利用 GeneMANIA 数据库对 *COL5A2* 参与的基因互作调控网络进行分析,结果以 *COL5A2* 基因为中心,呈扇形与多种基因互作

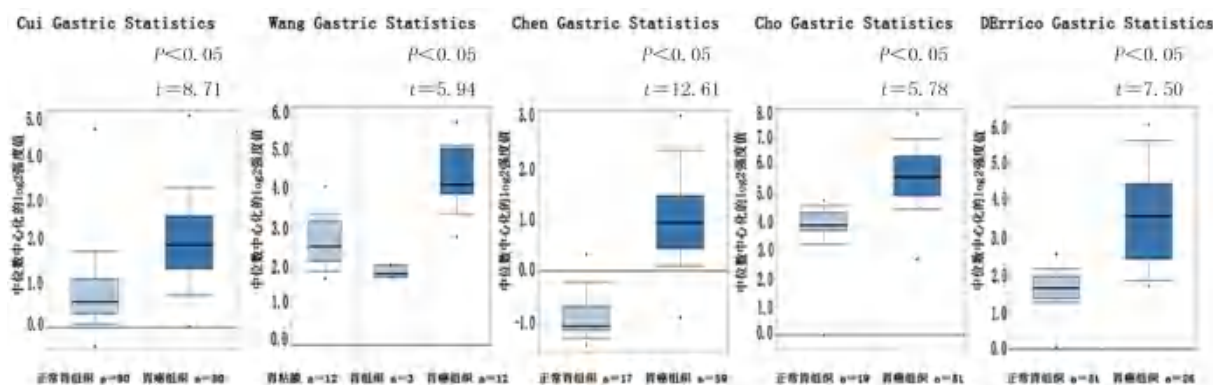
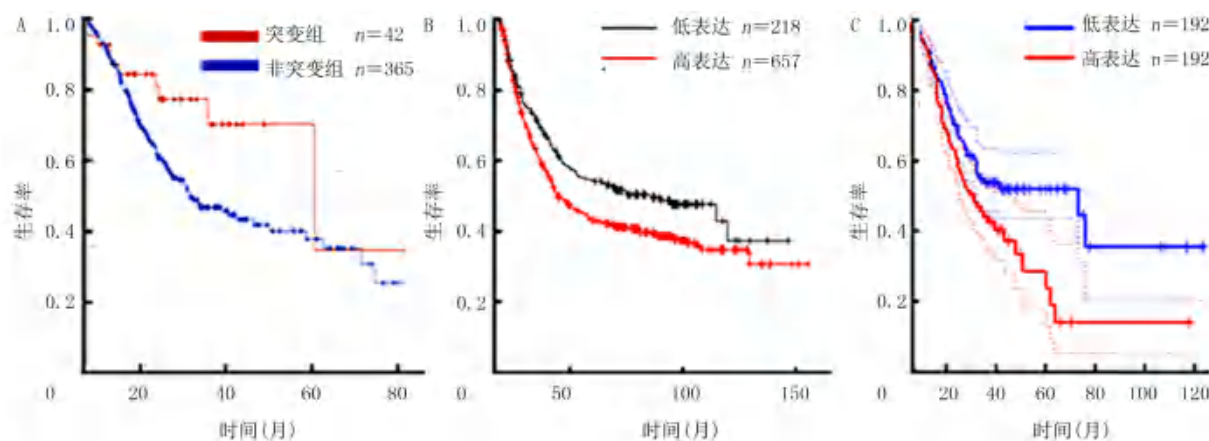


图 1 Oncomine 数据库 5 个 GC 芯片的 *COL5A2* 基因表达柱状图



A: cBioPortal 数据库生存差异分析图, B: Kaplan-Meier Plotter 数据库生存曲线图, C: GEPIA 数据库生存曲线图

图 2 不同数据库中 *COL5A2* 基因与 GC 患者的生存曲线图

展开,基因节点的大小表明了相互作用的强度,与 COL5A2 相互作用的基因互作的强度由强到弱分别包括 COL5A1、BMP1、ITGA1、COL1A1、COL3A1、COL1A2、COL6A3、TGM2、POSTN、SPARC、FN1、FBXW11、GP6、COL11A1、COL4A2、LAMC1、LUM、THBS2、MXRA5、FAP。见图 3。随后对其进行功能富集预测,结果显示相关基因主要与整合素通路、细胞外基质受体交互、细胞基质黏附、骨骼系统开发、肽交联、血管形态发生、细胞外基质分解及组织成形等有关。

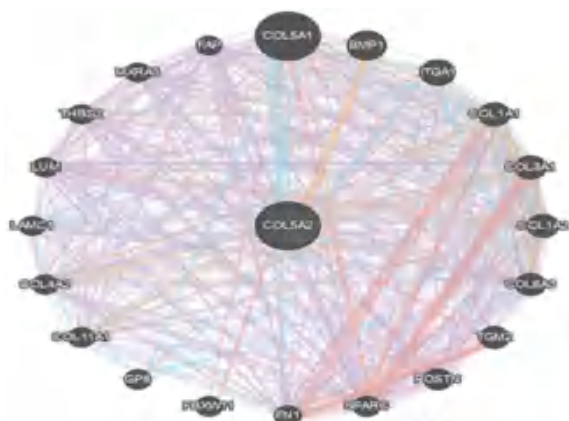


图 3 GeneMANIA 数据库基因互作网络图

## 2.6 GC 细胞系中 COL5A2 mRNA 的表达情况

qRT-PCR 结果显示, GES-1、HGC-27、MKN-45、BGC-823 和 SGC-7901 中 COL5A2 mRNA 相对表达量分别为  $1.00 \pm 0.00$ 、 $419.00 \pm 2.08$ 、 $228.00 \pm 1.53$ 、 $2.67 \pm 0.32$ 、 $3.87 \pm 0.19$ , 4 种 GC 细胞株 HGC-27、MKN-45、BGC-823 和 SGC-7901 COL5A2 mRNA 的表达水平均高于 GES-1, 差异具有统计学意义 ( $F = 253.40, P < 0.05$ ), 且在 MKN-45 和 HGC-27 细胞株中 COL5A2 mRNA 的表达量更高。因此, 进行后续细胞功能实验, 可选用 HGC-27 或者 MKN-45 细胞株。

## 2.7 COL5A2 siRNA 的转染效率

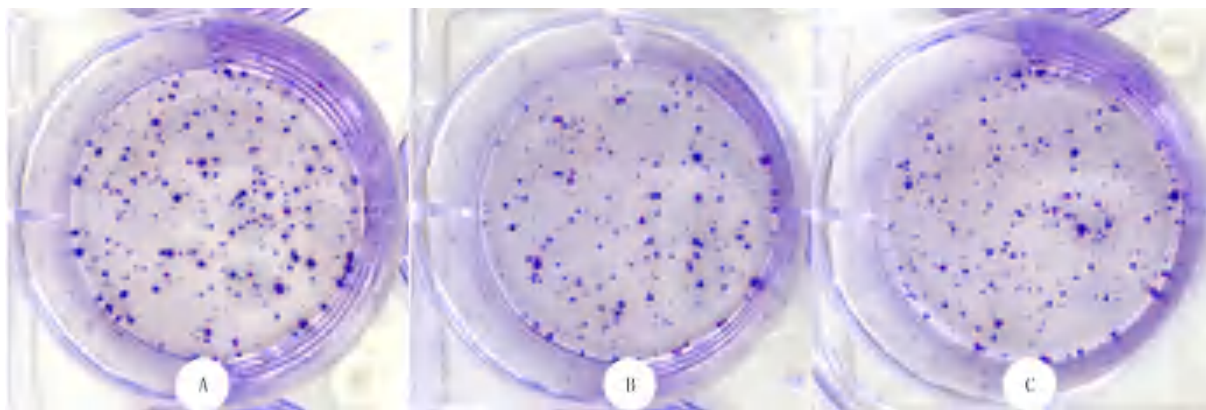
将 COL5A2 siRNA 转染 HGC-27 细胞, 培养 48 h 后采用 qRT-PCR 技术检测 COL5A2 基因干扰率。结果显示, 在 HGC-27 细胞中, si-1、si-2 和 si-3 转染组中 COL5A2 mRNA 的表达水平分别为  $0.23 \pm 0.04$ 、 $0.38 \pm 0.05$ 、 $0.82 \pm 0.04$ , 明显低于 si-NC 组 ( $1.00 \pm 0.00$ ) 的表达水平, 差异具有显著性 ( $F = 93.06, P < 0.05$ ), 但 si-1 和 si-2 干扰效果更为显著, 因此, 后续选用 si-1 和 si-2 进行细胞功能实验。

## 2.8 下调 COL5A2 基因后 HGC-27 细胞的增殖活力情况

CCK-8 实验结果显示, 时间、分组以及时间与分组交互作用对 HGC-27 细胞增殖活力均有显著的影响 ( $F = 4.50 \sim 223.60, P < 0.05$ )。3 组细胞转染 48、72 h 时的增殖活力与转染 0、24 h 比具有明显差异 ( $F = 38.90 \sim 70.30, q = 8.18 \sim 25.21, P < 0.05$ )。而与相应的 si-NC 组比较, 转染 72 h 时, si-1 和 si-2 组细胞增殖活力明显降低, 差异有统计学意义 ( $F = 21.29, q = 8.58, 9.87, P < 0.05$ ); 转染 0、24、48 h 时, si-1 和 si-2 组细胞增殖活力与相应的 si-NC 组比较无显著差异。见表 3。此外, 细胞克隆实验检测结果显示, si-1、si-2 和 si-NC 组形成的克隆数量分别为  $61.33 \pm 6.89$ 、 $72.33 \pm 8.35$ 、 $114.00 \pm 9.54$ , si-1 和 si-2 组形成的克隆数量明显少于 si-NC 组, 差异有统计学意义 ( $F = 11.12, P < 0.05$ )。见图 4。

表 3 不同转染时间对 HGC-27 细胞增殖活力的影响 ( $\bar{x} \pm s$ )

组别	0 h	24 h	48 h	72 h
si-NC 组	$46.10 \pm 5.13$	$55.14 \pm 6.20$	$76.49 \pm 4.86$	$138.80 \pm 3.87$
si-1 转染组	$37.75 \pm 4.78$	$48.81 \pm 5.74$	$67.84 \pm 3.64$	$102.50 \pm 3.96$
si-2 转染组	$39.20 \pm 4.98$	$47.40 \pm 4.60$	$70.83 \pm 4.76$	$107.30 \pm 4.92$



A: si-NC 组, B: si-1 转染组, C: si-2 转染组

图 4 不同处理组细胞克隆形成结果



### 3 讨 论

GC 是一种发病率高、治疗手段有限、临床疗效差的恶性肿瘤,开发新的生物标志物来预测 GC 患者的临床预后非常重要。细胞外基质处于一种动态平衡状态,通过调节纤维蛋白和各种胶原沉积而不断重构,调控肿瘤形成过程中的各种细胞行为<sup>[20]</sup>。已被证实多种胶原蛋白与人类多种肿瘤的发生与发展密切相关,如 V 型胶原蛋白家族成员 COL5A2。COL5A2 是细胞外基质中一个重要组成部分,与 I 型胶原形成异型纤维,从而在肿瘤浸润转移的初始过程中发挥重要作用<sup>[4-5,21]</sup>。然而,目前比较缺乏 COL5A2 基因在 GC 中具体表达和作用的相关研究。因此,本研究基于数据库挖掘探讨 COL5A2 基因在 GC 中的作用,为后续预后生物标志物的筛选、靶向药物的开发等提供一定的基础。

COL5A2 基因与人类多种类型肿瘤的发生发展密切相关。WU 等<sup>[22]</sup>利用生物信息学方法发现 COL5A2 基因在正常结肠组织中不表达,但在结直肠癌肿瘤组织中表达;同时 PATRA 等<sup>[23]</sup>报道了 COL5A2 mRNA 在结直肠癌组织中高表达;以上结果均提示 COL5A2 基因在结直肠癌发生发展中发挥促癌作用。另一方面,有研究发现 NKX2-2 基因在骨肉瘤中可通过介导 COL5A2 基因导致转录下调,从而起到抑制肿瘤的作用<sup>[24]</sup>。在对卵巢癌细胞(SKOV-3)和卵巢表面上皮细胞来源的外泌体蛋白质和脂质体分析中发现,SKOV-3 来源的外泌体中 COL5A2 蛋白显著高表达,提示其在卵巢癌早期诊断中具有重要作用<sup>[25]</sup>。此外,COL5A2 基因也可被视为肿瘤的生物标志物,与肿瘤患者的预后密不可分。ZENG 等<sup>[26]</sup>研究报道 COL5A2 mRNA 在膀胱癌中高表达,与肿瘤侵袭性、T 分期、N 分期密切相关,并且 COL5A2 mRNA 高表达患者比低表达患者的临床结局更差、生存率更低。同样的,WENG 等<sup>[27]</sup>研究发现 COL5A2 基因高表达的肺癌患者生存率低于低表达患者,认为 COL5A2 基因是肺癌患者的预后不良因素。值得注意的是,有研究发现在舌鳞状细胞癌患者中,COL5A2 基因高表达患者比低表达患者的无病生存期更长,认为 COL5A2 基因是舌鳞状细胞癌预后有利因素<sup>[28]</sup>。因此,COL5A2 基因在肿瘤中的作用是复杂的,在不同的癌症中发挥着不同的作用,对于 COL5A2 基因与特定肿瘤的关系,应当特定研究。

随着基因测序技术的发展,各种大规模的癌症

基因组项目收集了多维度的癌症基因组学数据,这给癌症基因组数据的研究、分析和整合提供了丰富的资源。Oncomine 是世界上最大的癌基因芯片在线数据库,包含全面癌症微阵列信息,有利于发现新型肿瘤生物标记物或治疗靶点<sup>[14]</sup>。cBioPortal 是一个多维度分析癌症基因组学数据的网络资源,可视化呈现基因和癌症样本之间的遗传变化<sup>[15]</sup>。Kaplan-Meier Plotter 是一个包含基因表达数据和临床数据的公共网络平台,可评估 54 000 个基因对包括 GC 在内的 21 种肿瘤预后的影响<sup>[16]</sup>。GEPIA 数据库整合了 TCGA 和 GTEx 数据,提供肿瘤样本和正常样本的 RNA 测序数据,对肿瘤相关联的基因表达谱数据进行动态分析,从而深入挖掘新型肿瘤标记物<sup>[17]</sup>。

本研究基于对以上肿瘤芯片数据库的挖掘,分析 COL5A2 基因在 GC 中的表达及其与 GC 患者临床预后的关系。首先在 Oncomine 数据库中本研究成果显示 COL5A2 基因在 GC 组织中高表达;通过 cBioPortal 数据库检索到的 415 例 GC 样本信息,表明 COL5A2 基因突变组比非突变组患者 OS 更长;在 Kaplan-Meier Plotter 平台进行生存分析显示,高表达 COL5A2 mRNA 的 GC 患者比低表达者预后更差。同时,GEPIA 数据库分析表明,COL5A2 mRNA 不仅在 GC 组织中高表达,且还与 GC 患者的肿瘤分期相关。此外,GEPIA 数据库的生存分析也支持前序结果,表明高 COL5A2 mRNA 表达与 GC 患者的不良临床预后显著相关,提示 COL5A2 基因是 GC 预后的不利因素。之后,本研究同时在分子水平上进行了验证,检测不同类型 GC 细胞系中 COL5A2 mRNA 的表达情况,发现 COL5A2 在 HGC-27 以及 MKN-45 胃癌细胞株中呈显著增高,与 COL5A2 mRNA 在 GC 组织中高表达的结果一致,因此,认为 COL5A2 基因在胃癌组织中发挥促癌基因的作用。

有关研究表明,COL5A2 基因可能通过缺氧、凝血、根尖连接、血管生成和凋亡等途径促进肿瘤的进展<sup>[25]</sup>。此外,MENG 等<sup>[29]</sup>研究发现 COL5A2 基因与上皮间质转化和细胞基质血管化高度相关,这些过程与癌症的侵袭以及转移均密切相关,推测 COL5A2 基因的表达可能与肿瘤侵袭密切相关。这在一定程度上解释了 COL5A2 基因对肿瘤发生预测的能力。GeneMANIA 是一个包含基因互作和潜在功能的平台在线数据库,而 Metascape 是一个提供全面的基因列表资源和分析的高效工具,通过

以上 2 种数据库,本研究评估了 *COL5A2* 密切关联的基因,从而进一步分析其可能参与的生物学途径与功能。首先,在分子水平上本研究发现 *COL5A2* 与 *COL5A1*、*MP1*、*ITGA1*、*COL1A1*、*COL3A1*、*COL1A2*、*COL6A3*、*TGM2*、*POSTN*、*SPARC*、*FN1*、*FBXW11*、*GP6*、*COL11A1*、*COL4A2*、*LAMC1*、*LUM*、*THBS2*、*MXRA5*、*FAP* 等基因互作调控,功能富集发现主要与整合素通路、细胞外基质受体交互、细胞基质黏附、骨骼系统开发、肽交联、血管形态发生及组织成形等有关。在这基础上,本研究还验证了 *COL5A2* 基因对 GC 细胞增殖和克隆形成能力的确切作用。CCK-8 和细胞克隆实验表明,在 HGC-27 细胞下调 *COL5A2* mRNA 表达后,细胞的增殖能力和克隆形成能力均明显下降。本研究仍有尚未探索的方面,今后应进一步探寻 *COL5A2* 基因在 GC 中的关键信号通路和下游靶基因等。

综上所述,*COL5A2* 基因在 GC 组织和 GC 细胞中高表达,且与 GC 患者的肿瘤分期和不良预后密切相关,*COL5A2* 基因突变患者预后相对较好,且 *COL5A2* 基因与细胞外基质受体交互、骨骼系统开发、肽交联等有关,下调 *COL5A2* 基因的表达可抑制 GC 细胞的增殖和克隆形成能力。*COL5A2* 基因在 GC 进展中起到重要作用,有望成为 GC 治疗的潜在新靶点。

#### [参考文献]

- [1] BARY F, FERLAY J, SOERJOMATARAM I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2018,68(6):394-424.
- [2] CHEN W Q, ZHENG R S, BAADE P D, et al. Cancer statistics in China, 2015 [J]. CA Cancer J Clin, 2016,66(2):115-132.
- [3] HASHIM D, BOFFETTA P, LA VECCHIA C, et al. The global decrease in cancer mortality: Trends and disparities[J]. Ann Oncol, 2016,27(5):926-933.
- [4] ISHIKAWA Y, TAGA Y, ZIENTEK K, et al. Type I and type V procollagen triple helix use different subsets of the molecular ensemble for lysine post-translational modifications in the rER[J]. J Biol Chem, 2021:e100453. Doi:10.1016/j.jbc.2021.100453.
- [5] ZHOU B, TU T, GAO Z, et al. Impaired collagen fibril assembly in keloids with enhanced expression of lumican and collagen V[J]. Arch Biochem Biophys, 2021,697:108676. Doi: 10.1016/j.abb.2020.108676.
- [6] MAK K M, PNG C Y M, LEE D J. Type V collagen in health, disease, and fibrosis[J]. Anat Rec, 2016,299(5):613-629.
- [7] THEOCHARIS A D, SKANDALIS S S, GIALELI C, et al. Extracellular matrix structure[J]. Adv Drug Deliv Rev, 2016, 97:4-27.
- [8] LI S, LIU X P, LIU T Z, et al. Identification of biomarkers correlated with the TNM staging and overall survival of patients with bladder cancer[J]. Front Physiol, 2017,8:947.
- [9] SONG Y X, WANG X Y, WANG F J, et al. Identification of four genes and biological characteristics of esophageal squamous cell carcinoma by integrated bioinformatics analysis[J]. Cancer Cell Int, 2021,21(1):123.
- [10] WEI S X, CHEN J S, HUANG Y, et al. Identification of hub genes and construction of transcriptional regulatory network for the progression of colon adenocarcinoma hub genes and TF regulatory network of colon adenocarcinoma[J]. J Cell Physiol, 2020,235(3):2037-2048.
- [11] ZHANG J F, SUN H, LIU S S, et al. Alteration of tumor-associated macrophage subtypes mediated by KRT6A in pancreatic ductal adenocarcinoma[J]. Aging, 2020,12(22):23217-23232.
- [12] WANG Y. Transcriptional regulatory network analysis for gastric cancer based on mRNA microarray[J]. Pathol Oncol Res, 2017,23(4):785-791.
- [13] WEI S Y, LU J H, LOU J Y, et al. Gastric cancer tumor microenvironment characterization reveals stromal-related gene signatures associated with macrophage infiltration [J]. Front Genet, 2020,11:663.
- [14] RHODES D R, KALYANA-SUNDARAM S, MAHAVISNO V, et al. OncoPrint 3.0: Genes, pathways, and networks in a collection of 18, 000 cancer gene expression profiles[J]. Neoplasia, 2007,9(2):166-180.
- [15] GAO J J, AKSOY B A, DOGRUSOZ U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal[J]. Sci Signal, 2013,6(269):p11.
- [16] GYÖRFFY B, LANCZKY A, EKLUND A C, et al. An online survival analysis tool to rapidly assess the effect of 22, 277 genes on breast cancer prognosis using microarray data of 1 809 patients[J]. Breast Cancer Res Treat, 2010,123(3): 725-731.
- [17] TANG Z F, LI C W, KANG B X, et al. GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses[J]. Nucleic Acids Res, 2017, 45 (W1): W98-W102.
- [18] FRANZ M, RODRIGUEZ H, LOPES C, et al. GeneMANIA update 2018[J]. Nucleic Acids Res, 2018,46(W1):W60-W64.
- [19] ZHOU Y Y, ZHOU B, PACHE L, et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets[J]. Nat Commun, 2019,10(1):1523.
- [20] ZHOU F F, SHEN D X, XIONG Y Y, et al. CTHRC1 is a prognostic biomarker and correlated with immune infiltrates in kidney renal papillary cell carcinoma and kidney renal clear cell

- carcinoma[J]. *Front Oncol*, 2020,10:570819.
- [21] LEE J S, MITULOVIĆ G, PANAHIPOUR L, et al. Proteomic analysis of porcine-derived collagen membrane and matrix[J]. *Materials (Basel)*, 2020,13(22):5187.
- [22] WU Y B, XU Y. Integrated bioinformatics analysis of expression and gene regulation network of COL12A1 in colorectal cancer[J]. *Cancer Med*, 2020,9(13):4743-4755.
- [23] PATRA R, DAS N C, MUKHERJEE S. Exploring the differential expression and prognostic significance of the COL11A1 gene in human colorectal carcinoma: An integrated bioinformatics approach[J]. *Front Genet*, 2021,12:608313.
- [24] CHEN H M, LIU W Q, ZHONG L, et al. NKX2-2 suppresses osteosarcoma metastasis and proliferation by downregulating multiple target genes[J]. *J Cancer*, 2018,9(17):3067-3077.
- [25] CHENG L, ZHANG K, QING Y N, et al. Proteomic and lipidomic analysis of exosomes derived from ovarian cancer cells and ovarian surface epithelial cells[J]. *J Ovarian Res*, 2020,13(1):9.
- [26] ZENG X T, LIU X P, LIU T Z, et al. The clinical significance of COL5A2 in patients with bladder cancer: A retrospective analysis of bladder cancer gene expression data[J]. *Medicine*, 2018,97(10):e0091.
- [27] WENG T Y, WANG C Y, HUNG Y H, et al. Differential expression pattern of THBS<sub>1</sub> and THBS<sub>2</sub> in lung cancer: Clinical outcome and a systematic analysis of microarray databases[J]. *PLoS One*, 2016,11(8):e0161007.
- [28] CHEN H C, TSENG Y K, SHU C W, et al. Differential clinical significance of COL5A1 and COL5A2 in tongue squamous cell carcinoma[J]. *J Oral Pathol Med*, 2019,48(6):468-476.
- [29] MENG X Y, SHI M J, ZENG Z H, et al. The role of COL5A2 in patients with muscle-invasive bladder cancer: A bioinformatics analysis of public datasets involving 787 subjects and 29 cell lines[J]. *Front Oncol*, 2018,8:659.
- (本文编辑 耿波 厉建强)

(上接第 5 页)

- [19] 中华医学会神经病学分会帕金森病及运动障碍学组. 中国帕金森病治疗指南(第三版)[J]. *药学与临床研究*, 2014,4:290.
- [20] KULISEVSKY J, OLIVEIRA L, FOX S H. Update in therapeutic strategies for Parkinson's disease[J]. *Curr Opin Neurol*, 2018,31(4):439-447.
- [21] LETA V, JENNER P, CHAUDHURI K R. Can therapeutic strategies prevent and manage dyskinesia in Parkinson's disease? [J]. *Expert Opinion on Drug Safety*, 2019,18(12):1203-1218.
- [22] MARRAS C, CHAUDHURI K R. Nonmotor features of Parkinson's disease subtypes[J]. *Mov Disord*, 2016,31(8):1095-1102.
- [23] WILSON H, GIORDANO B, TURKHEIMER F E, et al. Serotonergic dysregulation is linked to sleep problems in Parkinson's disease[J]. *Neuroimage Clin*, 2018,18:630-637.
- [24] XU D, ZHU W, HUO L, et al. Validation of Iodine-131-meta-iodobenzylguanidine cardiac scintigraphy in Parkinsonism: A preliminary study[J]. *Parkinsonism Relat Disord*, 2018,50:69-73.
- [25] MISHINA M, ISHII K, KIMURA Y, et al. Adenosine A1 receptors measured with <sup>11</sup>C-MPDX PET in early Parkinson's disease[J]. *Synapse*, 2017,71(8).Doi:10.1002/syn.21979.
- [26] CHEN S D, CHAN P, SUN S G, et al. The recommendations of Chinese Parkinson's disease and movement disorder society consensus on therapeutic management of Parkinson's disease[J]. *Transl Neurodegener*, 2016,5:12.
- [27] FOSSEL M. Use of telomere length as a biomarker for aging and age-related disease[J]. *Curr Transl Geriatr Exp Gerontol Rep*, 2012,1(2):121-127.
- [28] ERCAN A, KOHRT W M, CUI J, et al. Estrogens regulate glycosylation of IgG in women and men[J]. *JCI Insight*, 2017,2(4):e89703.
- [29] PAGANO G, NICCOLINI F, POLITIS M. Imaging in Parkinson's disease[J]. *Clin Med Lond Engl*, 2016,16(4):371-375.
- [30] RAMNANI D, KAPOOR S, UPRETI R. Unusual presentation of hyperhomocysteinaemia[J]. *Clinical Medicine*, 2019,19(Suppl 3):35.
- [31] KALILANI L, ASGHARNEJAD M, PALOKANGAS T, et al. Comparing the incidence of falls/fractures in Parkinson's disease patients in the US population[J]. *PLoS One*, 2016,11(9):e0161689.
- [32] PAUL S S, HARVEY L, CANNING C G, et al. Fall-related hospitalization in people with Parkinson's disease[J]. *Eur J Neurol*, 2017,24(3):523-529.
- [33] HINKLE J T, PEREPEZKO K, ROSENTHAL L S, et al. Markers of impaired motor and cognitive volition in Parkinson's disease: Correlates of dopamine dysregulation syndrome, impulse control disorder, and dyskinesias[J]. *Park Relat Disord*, 2018,47:50-56.
- [34] DRESSER R. Personal knowledge and study participation[J]. *J Med Ethics*, 2014,40(7):471-474.
- [35] CHEN S D, GAO G D, FENG T, et al. Chinese expert consensus on programming deep brain stimulation for patients with Parkinson's disease[J]. *Transl Neurodegener*, 2018,7:11.
- [36] 中国帕金森病脑深部电刺激疗法专家组. 中国帕金森病脑深部电刺激疗法专家共识[J]. *中华神经科杂志*, 2012,28(8):855-857.
- [37] KURTIS M M, RAJAH T, DELGADO L F, et al. The effect of deep brain stimulation on the non-motor symptoms of Parkinson's disease: A critical review of the current evidence[J]. *NPJ Parkinsons Dis*, 2017,3:16024.
- [38] LORING J F. Autologous induced pluripotent stem cell-derived neurons to treat parkinson's disease[J]. *Stem Cells Dev*, 2018,27(14):958-959.
- (本文编辑 耿波 厉建强)