81996_Auto_Edited.docx

WORD COUNT	4268	TIME SUBMITTED	04-APR-2023 03:55PM
		PAPER ID	98343151

Name of Journal: *World Journal of Gastrointestinal Oncology*

Manuscript NO: 81996

Manuscript Type: ORIGINAL ARTICLE

Basic Study

ENTPD1-AS1-miR-144-3p mediated high expression of COL5A2 correlates with poor prognosis and macrophage infiltration in gastric cancer

COL5A2 gastric cancer

Hanmei Yuan, Xiaofeng Pu, Hui Wu, Chao Wu

Abstract

BACKGROUND

Gastric cancer (GC) is a malignant tumor with high morbidity and mortality. Expression of *COL5A2* is significantly elevated in GC. Abnormal expression of ncRNAs has been found in GC, including miRNA and lncRNA. Competing endogenous RNA (ceRNA) network plays an important regulatory role in GC. However, its specific regulatory mechanism has not been elucidated.

AIM

To gain insight into the ncRNA regulatory mechanism and immune microenvironment related to *COL5A2* in GC.

METHODS

RNA sequencing data and clinical information from The Cancer Genome Atlas （TCGA）data portal were used to analyze the expressions of *COL5A2*, miRNA and lncRNA related to the prognosis of GC. Cox regression analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were performed to assess the risk factors and relevant function of *COL5A2* . StarBase was used to predict the interaction of miRNA–lncRNA or miRNA–mRNA in GC. The relationship between *COL5A2*, miR-144-3p and ENTPD1-AS1 was verified by dual luciferase reporter assay. The association of *COL5A2* with immune cell infiltration was analyzed using the Tumor Immune Estimation Resource (TIMER) database and single sample gene set enrichment analysis. The expression of *COL5A2* and macrophages in paired GC tissues were detected by immunohistochemical staining.

RESULTS

We verified that the upregulation of *COL5A2* expression was associated with the prognosis of GC and was an independent risk factor for GC. miR-144-3p was downregulated and correlated with the prognosis of GC. miR-144-3p regulated the

expression of *COL5A2* through direct interaction with miR-144-3p. ENTPD1-AS1 was elevated in GC and competitively bound to miR-144-3p, thus inhibiting the expression of miR-144-3p. ENTPD1-AS1 enhanced the expression of *COL5A2* through sponging miR-144-3p. Compared to paired normal tissue, *COL5A2* expression was upregulated at the protein level, especially in the middle and late stages of GC. The high expression of *COL5A2* was positively linked to macrophage infiltration in GC.

CONCLUSION

COL5A2 regulated by ENTPD1-AS1–miR-144-3p is associated with poor prognosis and macrophage infiltration in GC. This could be a new biomarker and therapeutic target in GC.

Key Words: *COL5A2*; ncRNAs; Macrophage infiltration; Prognosis; Gastric cancer

Yuan H, Pu X, Wu H, Wu C. ENTPD1-AS1-miR-144-3p mediated high expression of COL5A2 correlates with poor prognosis and macrophage infiltration in gastric cancer. *World J Gastrointest Oncol* 2023; In press

Core Tip: Gastric cancer(GC) is a malignant tumor with high fatality rate. ceRNA network and infiltration of immune cells play an important role in the development of GC. In this study, we verified that high expression of *COL5A2* was closely related to poor prognosis and was an independent risk factor for GC. We predicted and validated that lncRNA ENTPD1-AS1 regulated the expression of *COL5A2* through sponging miR-144-3p. Additionally, we confirmed that upregulation of *COL5A2* expression strongly correlated with immune infiltration of macrophages. ENTPD1-AS1-miR-144-3p-*COL5A2* might be a new therapeutic target for GC.

INTRODUCTION

Gastric cancer (GC) remains the fourth leading cause of cancer death worldwide. About 95% of cases are gastric adenocarcinoma (GAD) subtype[1]. The survival rate of GC has improved with medical advances. However, the median survival time for advanced GC is < 12 mo[2]. Accordingly, exploring the molecular mechanism of gastric cancer is of great significance for finding better therapeutic targets and better treatment.

The extracellular matrix is mainly composed of collagen and is related to proliferation, differentiation, migration, and metabolism of cancer[3]. Recent studies have shown that Collagen type V α2 (*COL5A2*), one of the collagen genes, is upregulated in several types of cancer and associated with immune cells infiltration[4, 5]. Traditionally, *COL5A2* is considered tightly related to the occurrence of classical Ehlers–Danlos syndrome [6, 7]. However, high expression of *COL5A2* is found to be associated with worse prognosis and drug resistance[8-10]. *COL5A* is also found closely associated with immune cell infiltration, which may be related to the inhibitory effect of collagen on the production of CCL2[5]. In proliferative diabetic retinopathy, *COL5A2* is closely related to the infiltration of M2 macrophages[11].

Noncoding RNAs (ncRNAs), such as microRNAs (miRNAs), long noncoding RNAs (lncRNAs) and circular RNAs (circRNAs), are critically involved in GC development[12]. ncRNAs could affect the proliferation, invasion, migration and metastasis of GC by regulating target miRNA genes[13]. Studies have revealed that MiR-144-3p can affect the occurrence, development, and prognosis of cancer by inhibiting the expression of target genes[14, 15]. ncRNA can upregulate the expression of cancer-promoting genes by inhibiting the expression of miRNA-144-3p[16, 17]. ENTPD1-AS1 is an antisense lncRNA that may be associated with short stature[18]. Recently, ENTPD1-AS1 is considered to be a new ncRNA that regulates the proliferation and apoptosis of cancer cells and also serves as a prognostic marker in glioblastoma multiforme[19, 20]. However, the ncRNA regulatory mechanism and immune infiltration of *COL5A2* are still unclear in GC.

In this study, we verified that high expression of *COL5A2* in GC was closely related to poor prognosis and was an independent risk factor. We predicted and verified a new

upstream regulatory ceRNA network, namely, ENTPD1-AS1 regulated *COL5A2* expression through sponging miR-144-3p. We confirmed that upregulation of *COL5A2* expression strongly correlated with immune infiltration of macrophages. ENTPD1-AS1–miR-144-3p regulation of *COL5A2* correlated with poor prognosis and macrophage infiltration in GC.

MATERIALS AND METHODS

Sample collection

We collected 40 paired GC and normal specimens from the Eighth Affiliated Hospital of Sun Yat-Sen University. These tissues were obtained from the patients undergoing GC surgery. All the patients signed informed consent forms. This research was approved by the Ethics Committee of the Eighth Affiliated Hospital of the Sun Yat-Sen University.

Cell culture

Human Embryonic Kidney cells (293T) and human gastric adenocarcinoma cell line (AGS) were cultured in dulbecco's modified eagle medium (DMEM) (GIBCO, Invitrogen, Carlsbad, CA, USA) containing 10% FBS (GIBCO) and 100 U/mL penicillin, and 0.1 mg/mL streptomycin (Invitrogen, Carlsbad, CA, USA).

Expression and prognostic analysis of COL5A2 in gastric cancer

The Gene Expression Profiling Interactive Analysis (GEPIA; http://gepia.cancer-pku.cn/) database was utilized for the expression and prognostic analysis of *COL5A2* in numerous cancers.

Clinical information and data acquisition of patients with GAD

The mRNA expression data of GAD in tumor and normal tissues (446 cases of tumor, and 45 cases of normal tissue) and clinical material were obtained from TCGA public database (https://portal. gdc. cancer. gov). The basic information of human miRNA was downloaded from http://www.mirbase.org. We conducted a differential analysis and survival analysis on the expression of *COL5A2* in normal tissues and gastric cancer

tissues. Data transformation was achieved using Perl script. R version 4.1.2 was used in several analyses.

Cox regression analysis

We organized the clinical data of GAD and matched the expression of *COL5A2* in the corresponding samples. Samples with incomplete or missing data were excluded from this analysis, and finally 322 specimens were obtained. Univariate and multivariate Cox proportional hazards regression models were used to assess *COL5A2* for prediction of overall survival (OS).

Functional enrichment analysis of COL5A2

We analyzed Kyoto Encyclopedia of Genes and Genomes (KEGG) signaling pathway enrichment of *COL5A2* by gene set enrichment analysis (GSEA). We found gene markers corresponding to 22 immune cells and extracted the expression of these gene markers from the TCGA data. Using single sample GSEA (ssGSEA), we determined the infiltration of these immune cells in *COL5A2* high expression and low expression groups.

Prediction and analysis of the interaction between mRNA and miRNA or miRNA and lncRNA

We used starBase (http://starbase.sysu.edu.cn/) to predict the interaction of miRNA–mRNA or miRNA–lncRNA interaction. StarBase is a powerful target gene prediction software, and it includes 7 functional domains for prediction. We set the screening conditions for at least two sites that considered the gene to be the target gene of *COL5A2*, and the gene was retained. The expression data of miRNA-144-3p and ENTPD1-AS1 are from the TCGA database. We conducted the expression analysis, correlation analysis, and survival analysis using R version 4.1.2.

Dual-luciferase reporter assay

The sequences of *COL5A2* or ENTPD1-AS1 that may bind to miRNA-144-3p were cloned into the pmirGLO vector (Sangon Biotech Co., Ltd. Shanghai, China). We constructed four kinds of double luciferase report plasmids: COL5A2 3'UTR-wt, COL5A2 3'UTR -mut, ENTPD1-AS1-wt and ENTPD1-AS1-mut.

SiRNAs and miRNA mimics

siRNAs,miRNA mimics and their corresponding negative controls (NC) were designed by GenePharma Co., Ltd., (Shanghai, China). The sequence is as follows: MiR-144-3p mimics:(UACAGUAUAGAUGAUGUACU), mimics NC: (UUGUACUACACAAAAGUACUG), si-ENTPD1-AS1:(GGCCCGUAAUGGAGAUCGATT,UCGAUCUCCAUUACGGGCCTT),si-NC:(UUCUCCGAACGUGUCACGUTT, ACGUGACACGUUCGGAGAATT).

Cell transfection

We seeded cells into 24-well plates at a density of $1.5x10^5$ cells per well. When the cells reached 60% to 70%, dual luciferase reporter vector and miR-144-3p mimics or negative control (NC) mimics were transfection into 293T cells and AGS cells, in the presence of Lipofectamine™ and P3000™ (L3000001, Invitrogen, USA). In some experiments, mixture of Lipofectamine™ and miR-144-3p or si-ENTPD1-AS1 was transfection into AGS cells. After 6 h, fresh medium was replaced. 48h later, Fluorescence intensity is detected by Dual-Luciferase® Reporter Assay System ((E1910, Promega, USA).

qRT-PCR

RNAs were extracted by Trizol from AGS cells. RNA was reversely transcribed into cDNA with Evo M-MLV RT Premix (AG11701, Accurate Biotechnology, Hunan, China.) and then detected gene expression with SYBR® Green Premix Pro Taq HS qPCR Kit for qPCR (AG1170, Accurate Biotechnology, Hunan, China). The primers were synthesized by Sangon Biotech Co., Ltd (Shanghai, China). *COL5A2*: (Forward: GGATCACAGGGACCAAGAGGAGAG, Reverse:

GCACCAGGTTGACCAGGAACAC), ENTPD1-AS1(Forward: CCTGCCTCTGCCTCCAAGTAG, Reverse: TTCGAGACCAGCCTGACCAAC), hsa-miR-144-3p (RT Primer: GTCGTATCCAGTGCAGGGTCCGAGGTATTCGCACTGGATACGACAGTACA, Forward: GCGCGCGTACAGTATAGATGA, Reverse: ATCCAGTGCAGGGTCCGAGG). U6 (Forward: GGAACGATACAGAGAAGATTAGC, Reverse: TGGAACGCTTCACGAATTTGCG), GAPDH (Forward: TGTGTCCGTCGTGGATCTGA Reverse: GCAGCTGTGACACACAGTA). miRNA was detected by stem-loop. U6 or GAPDH as internal control. The relative expression of genes is calculated by $2^{-\Delta\Delta Ct}$.

Correlation Analysis of COL5A2 and Immunological Checkpoints

Tumor IMmune Estimation Resource (TIMER) database (https://cistrome.shinyapps.io/timer/) integrated the results of multiple algorithms based on the sequencing results of TCGA. We used the TIMER database and GEPIA to analyze the correlation between *COL5A2* and immune cells or immune checkpoints.

Immunohistochemistry

All tissue slides were dewaxed, hydrated, and antigen was retrieved. We inactivated endogenous enzymatic activity and blocked nonspecific sites. Primary and secondary antibodies were incubated and rendered with Diaminobenzidine(DAB）. The concentration of primary antibodies was anti-*COL5A2* (1:100, Thermo Fisher, PA5-38880, USA), anti-CD68 (1:2000, Abcam, ab955, UK). We selected the corresponding secondary antibody according to the primary antibody (1:100, Santa Cruz Biotechnology, sc-2357/sc-516102; DAB detection kit (50:1, Servicebio, G1212, China). All slides were observed and counted by a Carl Zeiss microscope (Axio Imager A2, Germany). The $CD68^+$ cell counts were conducted by taking the average value of three high power fields (HP). Based on the number of $CD68^+$ cells, the macrophage infiltration was categorized into weak (30–60/HP), moderate (60–90/HP), and strong (> 90/HP) subgroups.

Statistical analysis

R with the survminer, survival, Ggforest and limma packages was used for analysis. GraphPad 8 with unpaired *t*-test and Spearman correlations were used for some analyses. $P < 0.05$ was considered statistically significant.

RESULTS

Expression and prognostic analysis of COL5A2 in pan-cancer

To identify the differential expression of *COL5A2* in pan-cancer, the expression of *COL5A2* was assessed in 18 types of cancer by GEPIA. The expression of *COL5A2* was significantly increased in nine cancers compared with the normal group (Figure 1A–1I, $p<0.05$). There was no significant difference in the other nine cancers (Supplementary Figure 1). Among the cancers with significant differences in *COL5A2* expression, OS analysis was conducted using GEPIA database. Only in GC was higher expression of *COL5A2* related to worse prognosis (Figure 1Q, $P = 0.01$). No significant prognostic difference in other cancers was observed (Figure 1J–1R).

Cox regression analysis and KEGG enrichment analysis of COL5A2 in GC

To verify the results obtained from the GEPIA database, we analyzed the expression and survival prognosis of *COL5A2* using RNA-sequencing data from TCGA. In line with previous studies[21], significantly high expression of *COL5A2* was found in TCGA data and 27 paired specimens (Figures 2A and 2B). We performed survival analysis and found a significant difference in OS between high and low COL5A2 expression groups (Figure 2C, $P = 0.018$).

Excluding patients with missing or incomplete information, 322 of the 433 patients were included in the cox regression analysis. From clinical information, it can be seen that patients with gastric cancer are mainly in the middle and late stages, with a high rate of lymph node metastasis, less distant metastasis, and a high mortality rate (Supplementary Table 1). Cancer stage, especially N stage, was the most important

prognostic indicator. Age, T stage, and *COL5A2* expression were risk factors for GC (Table 1). To determine whether these risk factors were independent of other factors as independent prognostic factors, multivariate Cox regression was conducted. Only age [1.03(1.01-1,1), P = 0.002] and expression of *COL5A2* [1.23(1.05-1.5), P = 0.013] were independent prognostic factors (Figure 2D).

Dividing GCs into *COL5A2* high- and low-expression groups, we used KEGG enrichment to analyze which signaling pathways were associated with differential genes between the two groups. According to Normalized enrichment score (NES), nominal P value, the 10 most enriched signaling pathways in KEGG are shown in Supplementary Table 2). To be more concise, we integrated these signaling pathways into one diagram (Figure 2E). The most significantly enriched signaling pathways in the *COL5A2* high expression group were the interaction between extracellular matrix and receptors, focal adhesion, and some classic cancer-related signaling pathways such as the transforming growth factor-β signaling pathway.

miRNAs that interact with COL5A2 in GC

To investigate whether *COL5A2* was regulated by miRNAs in GC, we used starBase to predict which miRNAs would bind to *COL5A2*. There were 69 miRNAs with the potential to bind to *COL5A2*, and it was visualized by cytoscape (Supplementary Figure 2). According to the principle of gene regulation, miRNA should negatively correlate with *COL5A2*. Thus, we set the screening condition that miRNA was negatively correlated with *COL5A2* and $P < 0.001$. For *COL5A2*, there was upregulation in GC, so the target miRNA downregulation in GC compared with normal tissues. miR-29c-3p (R= -0.27, P = 2.1e-07) and miR-144-3p (R=-0.16, P = 0.0017) were negatively correlated with *COL5A2* and markedly downregulated in GC (Figures 3A, 3B (P = 1.4e-09), 3D and 3E (P = 0.013). We analyzed the prognostic impact of miR-29c-3p and miR-144-3p. Only low expression of miR-144-3p had a significant effect on survival (Figures 3C (P = 0.058) and 3F (P = 0.041). All these findings meant that miR-144-3p is the most likely upstream miRNA to regulate *COL5A2* in GC.

Prediction and analysis of upstream lncRNAs of miR-144-3p

lncRNAs can competitively bind to miRNAs, which leads to upregulation of oncogenes. We predicted using starBase which lncRNAs would interact with miR-144-3p. We examined the correlation between lncRNAs, *COL5A2* and miR-144-3p (Table 2). For downregulated miR-144-3p, expression of lncRNAs should be upregulated in GC. Among all lncRNAs, ENTPD1-AS1(R=-0.32, *P* = 1.6e-10), NORAD (R=-0.21, *P* = 3e-05), and ZNF460-AS1(R=-0.24, *P* = 2.1e-6) were negatively correlated with miR-144-3p (Figures 4A, 4E and 4I). ENTPD1-AS1(*r* = 0.11, *P* = 0.036) and NORAD (*r* = 0.14, *P* = 0.0067) were positively correlated with *COL5A2*. ZNF460-AS1(R=-0.042, *P* = 0.42) had a negative correlation with *COL5A2* (Figures 4B, 4F and 4J). We analyzed the expression of lncRNA in GC, and three lncRNAs were significantly upregulated in GC (Figures 4C(*P* = 3.6e-06), 4G(*P* = 1e-10) and 4K(*P* = 0.011)). Kaplan–Meier analysis revealed that the higher the ENTPD1-AS1 expression, the better the OS (Figures 4D(*P* = 0.002), 4H(*P* = 0.229) and 4L(*P* = 0.351). By taking into account expression and prognostic analysis, our data suggest that ENTPD1-AS1 is the most likely lncRNA to regulate the miR-144-3p/*COL5A2* axis in GC.

LncRNA ENTPD1-AS1 enhances the expression of COL5A2 through sponging miR-144-3p

In order to verify the interaction of ENTPD1-AS1, miR-144-3p and *COL5A2*, we constructed dual luciferase reporter vector (Figures 5A). The results showed that in the presence of miR-144-3p mimics, the luciferase activity of *COL5A2* 3'UTR-wt group was significantly reduced compared to *COL5A2* 3'UTR-mut group in both 293T cells and AGS cells. However, when transfected with negative control(NC) mimics, there was no difference in fluorescence intensity between *COL5A2* 3'UTR-wt group and *COL5A2* 3'UTR-mut group (Figures 5B and 5C). Similarly, the luciferase activity was reduced when miR-144-3p mimics interacted with ENTPD1-AS1-wt instead of ENTPD1-AS1-mut (Figures 5D and 5E). Additionally, when miR-144-3p enrichment was present, the

expression of *COL5A2* was decreased in AGS cells. When transfected with NC mimics, the change of fluorescence intensity disappeared (Figures 5F and 5G). Compared to si-NC, si-ENTPD1-AS1 Led to a decrease of miR-144-3p (Figures 5H and 5I). All these results suggested that ENTPD1-AS1 might promote the expression of COL5A2 by suppressing the expression of miR-144-3p.

COL5A2 has a correlation with immune checkpoints in GC

For immunotherapy having a low response in many cancers, there is a pressing need to find new target genes that could improve the efficacy of immunotherapy. Considering the cancer-promoting role of *COL5A2*, we analyzed the association of *COL5A2* with checkpoints using the TIMER database. Our analyses revealed a high correlation between *COL5A2* and CD274 and a lower correlation between *COL5A2* and CTLA4 or PDCD1 (Figure 6A). We observed similar results in the GEPIA database (Figures 6B–6D). These results suggest that *COL5A2* mediates immune escape in GC.

COL5A2 is associated with immune cell infiltration in GC

To explore the link between *COL5A2* and immune cells, we determined the relationship between *COL5A2* and immune cell biomarkers. We found a significant and positive association with *COL5A2* and M2 macrophage biomarkers (CD163, VSIG4 and MS4A4A). The other immune cells biomarkers had a lower correlation or only some of the markers had a correlation with *COL5A2* (Table 3). The TIMER database indicated that *COL5A2* was significantly associated with infiltration of various immune cells, including neutrophils, macrophages and dendritic cells, but most correlated with macrophage infiltration (Figure 7A). We used ssGSEA to analyze the enrichment of 22 types of immune cells in groups with high and low *COL5A2* expression. We discovered a significant difference in immune infiltration of macrophages, neutrophils and B memory cells between *COL5A2* high and low expression groups. In line with our previous results, macrophage infiltration was most correlated with *COL5A2* (Figure 7B, $p<0.001$). *COL5A2* is significantly positively linked to macrophage infiltration.

Expression of COL5A2 and its correlation with macrophages in GC

To verify the correlation between *COL5A2* and macrophage infiltration in GC, we detected expression of *COL5A2* and macrophages ($CD68^+$) by immunohistochemistry in GC and paired normal tissues. According to the stain density, we divided the expression of *COL5A2* into three levels, including weak, moderate, and strong staining, with corresponding scores of 1–3. According to the number of $CD68^+$ cells, we divided macrophage infiltration into three groups: weak (30–60/HP), moderate (60–90/HP), and strong (> 90/HP) (Figure 8A). The expression of *COL5A2* at the protein level was clearly higher than in the paired peritumor tissues (Figure 8B, $P = 0.025$). Expression of *COL5A2* between low stage GC (I and II) and high stage GC (III and IV) was assessed to determine whether *COL5A2* was related to GC stage. Compared with low stage GC, the level of *COL5A2* was significantly higher in high stage GC (Figure 8C, $P = 0.0106$). Pearson correlation analysis was performed to determine the relation between *COL5A2* expression and macrophage infiltration in GC. Clearly, macrophage infiltration was consistent with *COL5A2* staining intensity (Figure 8D, $r = 0.7459$, $p<0.0001$). These results demonstrated that high expression of *COL5A2* was observed in GC and *COL5A2* was strongly positively correlated with macrophage infiltration.

DISCUSSION

GC is one of the tumors with high morbidity and mortality. Collagen is upregulated in advanced GC, and some collagen genes can be used as cancer biomarkers and can distinguish precancerous from cancerous lesions[22, 23]. Several studies have shown that *COL5A2* is upregulated in various cancers and can be used as a prognostic marker [24, 25]. Here, we explored the upstream regulatory molecular mechanisms and immune function of *COL5A2* in GC. We confirmed that *COL5A2* was enriched and can predict poor prognosis in GC. We verified that *COL5A2* is closely related to macrophage infiltration in GC. A new ceRNA network, ENTPD1-AS1–miR-144-3p–*COL5A2* was

identified, which may partially explain the upstream regulatory mechanism of *COL5A2* in GC.

We confirmed that *COL5A2* was increased in a number of tumors, but it was associated with worse survival only in GC. Cox regression analysis found that *COL5A2* could be an independent prognostic factor for GC. KEGG enrichment analysis found that high expression of *COL5A2* was associated with multiple signal transduction pathways. All these results revealed that *COL5A2* was strongly associated with worse prognosis and has a crucial role in the development of GC. ceRNA plays a critical role in the regulation of gene expression in GC [26, 27]. The mechanism of ceRNA mainly refers to that lncRNAs inhibit the negative regulation of miRNAs. For example, upregulation of circular RNA KIF4A expression in GC can regulate the miRNA-144-3p–EZH2 axis to facilitate cell migration and invasion[28]. To investigate whether *COL5A2* was regulated by the ceRNA network, we predicted by starBase which miRNAs could bind to *COL5A2*. We discovered that miR-29c-3p and miR-144-3p were most likely to regulate *COL5A2* in GC. Although miR-29c was reduced in GC and associated with tumor aggressiveness, our findings suggest that miR-29c-3p plays a role in GC by regulating *COL5A2*. Some studies have demonstrated that miR-144-3p inhibited cancer proliferation and migration, potentially serving as a biomarker in GC [16, 29], which is consistent with our results that miR-144-3p was downregulated and closely related to prognosis in GC. Considering correlation analysis and expression analysis, ENTPD1-AS1, NORAD, and ZNF460–AS1 might regulate *COL5A2* through a ceRNA network in GC. Three lncRNAs were negatively correlated with miR-144-3p, but only ENTPD1-AS1 and NORAD were positively correlated with *COL5A2*. Therefore, we focused on ENTPD1-AS1 and NORAD. Survival analysis suggested that ENTPD1-AS1 was significantly associated with GC survival. The most potentially upregulated lncRNAs were ENTPD1-AS1. Many lncRNAs have been identified as ceRNAs in GC, such as HOTAIR, MALAT1, NORAD and H19[30]. Some studies reported that lncRNA NORAD promoted GC cell growth by inhibiting expression of miR-608 or miR-433-3p [31, 32]. Here, we revealed that NORAD promoted the development of GC by inhibiting

expression of miR-144-3p. ENTPD1-AS1 is an antisense transcription lncRNA, which can be used as a prognostic marker in glioblastoma multiforme [19]. We confirmed that the most probable ceRNA regulatory network for *COL5A2* in GC is NTPD1-AS1–miR-144-3p–*COL5A2* through dual luciferase assay.

Recent studies have shown that fibrillar collagen can facilitate immune cell infiltration with bioinformatics analysis [33]. Genes associated with M2 infiltration of GC have been described in some previous studies, including COL1A1, COL4A1, COL12A1 and PDGFRB [34, 35]. Wei *et al* [36]identified that stromal-relevant gene clusters could be used as prognostic genes and are associated with macrophage infiltration in GC. However, most of the results were obtained through database analysis and lacked experimental validation. Similar to previous studies, we found that *COL5A2* was markedly positively associated with macrophage infiltration, using the TIMER database and ssGSEA. We confirmed by immunohistochemical staining that *COL5A2* was significantly highly expressed, especially in the high stage of GC and was significantly positively correlated with macrophage infiltration at the protein level by IHC staining. These results give us a more complete view that macrophage infiltration may partially explain the carcinogenesis mediated by *COL5A2* in GC.

There are some limitations to this study. First, although we proved the direct interaction of miR-144-3p and *COL5A2* or miR-144-3p and ENTPD1-AS1 by Dual-Luciferase Reporter Assay and siRNA transfection, further experimental verification and confirmation are needed, such as RNA Binding Protein Immunoprecipitation (RIP) and *in vivo* experiments. Second, *COL5A2* was found to be associated with poor prognosis of GC and macrophage infiltration by *in vitro* experiments, the sample size was small, these results thus should be verified in *COL5A2* knockout mice, and tested in big samples.

CONCLUSION

We have verified that *COL5A2* is an independent risk factor and can be used as a biomarker for GC. Our results demonstrate that *COL5A2* exerts a tumor-promoting effect by promoting immune cell infiltration, especially macrophage infiltration. We have identified a novel ceRNA network that facilitates *COL5A2* expression in GC, namely, lncRNA ENTPD1-AS1 upregulates the expression of *COL5A2* by inhibiting the expression of miR-144-3p. These results partly explained the upstream regulatory mechanism and immune mechanism of *COL5A2* in GC. *COL5A2*–miR-144-3p–ENTPD1-AS1 has the potential to be a novel therapeutic target for GC.

ARTICLE HIGHLIGHTS

Research background

Gastric cancer (GC) is a malignant tumor with high morbidity and mortality. Expression of *COL5A2* is significantly elevated in GC. However, its specific regulatory mechanism has not been elucidated.

Research motivation

Abnormal expression of ncRNAs has been found in GC, including miRNA and lncRNA. The ncRNA regulatory mechanism and immune microenvironment related to *COL5A2* in GC are not well understand.

Research objectives

To explore the ceRNA regulatory mechanism and immune mechanism of *COL5A2* in gastric cancer.

Research methods

StarBase was used to predict the interaction of miRNA–lncRNA or miRNA–mRNA in GC. The direct interaction between *COL5A2*, miR-144-3p and ENTPD1-AS1 was verified by dual luciferase reporter assay. The correlation between of *COL5A2* and

macrophages was analyzed through bioinformatics and validated in paired GC tissues by immunohistochemical staining.

Research results

miR-144-3p interacted directly with *COL5A2* and negatively regulated the expression of *COL5A2* . ENTPD1-AS1 was elevated in GC and competitively bound to miR-144-3p, thus inhibiting the expression of miR-144-3p. Compared to paired normal tissue, *COL5A2* expression was upregulated at the protein level, especially in the middle and late stages of GC. The high expression of *COL5A2* was positively linked to macrophage infiltration in GC.

Research conclusions

COL5A2 regulated by ENTPD1-AS1–miR-144-3p is associated with poor prognosis and macrophage infiltration in GC.

Research perspectives

ENTPD1-AS1-miR-144-3p-*COL5A2* might be a new therapeutic target for GC.

81996_Auto_Edited.docx

ORIGINALITY REPORT

3%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.ncbi.nlm.nih.gov Internet	32 words — 1%
2	www.researchgate.net Internet	32 words — 1%
3	www.researchsquare.com Internet	18 words — < 1%
4	www.wjgnet.com Internet	18 words — < 1%
5	Xingqin Zhou, Yazhou Chang, Lirong Zhu, Chaoyan Shen, Jing Qian, Renan Chang. "LINC00839/miR-144-3p/WTAP (WT1 Associated protein) axis is involved in regulating hepatocellular carcinoma progression", Bioengineered, 2021 Crossref	14 words — < 1%
6	Butian Li, Shengping Zhang, Hao Shen, Chenglong Li. "MicroRNA-144-3p suppresses gastric cancer progression by inhibiting epithelial-to-mesenchymal transition through targeting PBX3", Biochemical and Biophysical Research Communications, 2017 Crossref	12 words — < 1%
7	Junjie Chen, Cui'e Lu, Xiaodong Wang, Lin Wang, Jia Chen, Fuhai Ji. "LncRNA NONRATT009773.2	12 words — < 1%

promotes bone cancer pain progression through the miR‐708‐5p/CXCL13 axis", European Journal of Neuroscience, 2022
Crossref

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	< 12 WORDS
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	< 12 WORDS