Long Noncoding RNA GAPLINC Regulates CD44-Dependent Cell Invasiveness and Associates with Poor Prognosis of Gastric Cancer

Ye Hu1, Jilin Wang1, Jin Qian2, Xuan Kong1, Jieting Tang1, Yingchao Wang1, Haoyan Chen2, Jie Hong2, Weiping Zou3, Yingxuan Chen1, Jie Xu1, and Jing-Yuan Fang1

Abstract

It is increasingly evident that long noncoding RNAs (lncRNA) have causative roles in carcinogenesis. In this study, we report findings implicating a novel lncRNA in gastric cancer, termed GAPLINC (gastric adenocarcinoma predictive long intergenic noncoding RNA), based on the use of global microarray and in situ hybridization (ISH) analyses to identify aberrantly expressed IncRNA in human gastric cancer specimens. GAPLINC is a 924-bp-long lncRNA that is highly expressed in gastric cancer tissues. GAPLINC suppression and with gene expression profiling in gastric cancer cells revealed alterations in cell migration pathways, with CD44 expression the most highly correlated. Manipulating GAPLINC expression altered CD44 mRNA abundance and the effects of GAPLINC on cell migration and proliferation were neutralized by suppressing CD44 expression. Mechanistic investigations revealed that GAPLINC regulates CD44 as a molecular decoy for miR211-3p, a microRNA that targets both CD44 and GAPLINC. Tissue ISH analysis suggested that GAPLINC overexpression defines a subgroup of patients with gastric cancer with very poor survival. Taken together, our results identify a noncoding regulatory pathway for the CD44 oncogene, shedding new light on the basis for gastric cancer cell invasiveness. Cancer Res; 74(23); 1–13. ©2014 AACR.

Introduction

Gastric cancer is the second leading cause of cancer-related mortality in the world, and the majority of patients with gastric cancer are diagnosed at an advanced stage and die within 24 months after operation because of recurrence and metastasis (1, 2). To improve gastric cancer early diagnosis and targeted therapy, an in-depth understanding of molecular underpinnings of the disease is required (3–5). It is of clinical importance to identify genes that contribute to gastric cancer development and present predictive value for diagnosis or prognosis (6–8). Currently, most reported potential biomarkers for gastric cancer are protein-coding genes (PCG), including the novel somatic gene targets (ARID1A, FAT4, MLL, and KMT2C) revealed by large-scale cancer genomic studies. Despite extensive efforts to develop PCG-based biomarkers, only modest successes have been obtained in biomarker-assisted gastric cancer diagnosis and treatment (3, 9).

Recent integrative genomic studies have revealed that the human genome encodes more than 10,000-long noncoding RNAs (lncRNA) with limited or no protein-coding capacity (10–13). Although a small number of lncRNAs have been functionally characterized (14–17), a large number of members in the class remain functionally uncharacterized (18, 19). Growing evidence suggests that cancer lncRNAs, similar to PCGs, may mediate oncogenic or tumor-suppressing effects and may be a new class of cancer biomarkers and therapeutic targets (20–22). One such lncRNA is HOTAIR, which is expressed from the developmental HOXC locus and associates with chromatin modifications in cooperation with the Polycomb complex PRC2 (23). Overexpression of HOTAIR is a powerful predictor of the tumor progression and overall survival in patients with diverse cancers, including diffuse type of gastric cancer (24). Recent studies also suggest that lncRNAs can also act as decoys for microRNAs, and an example of this mechanism is represented by the tumor-suppressor gene PTEN and its pseudogene PTENP1. The PTENP1 3′-untranslated region (3′-UTR) was found to increase PTEN expression by binding to microRNAs that downregulate PTEN expression. Despite the above findings, our current knowledge about the expression patterns and functional roles of lncRNAs in gastric cancer is still limited. In previous studies, efforts have been made to reannotate the probes of Affymetrix microarrays that match lncRNA sequences in several cancers, including gastric cancer (25). However, this method covers less than 65% of lncRNA genes, and thus should not be considered as unbiased analysis.
In this study, we identify deregulated lncRNAs in gastric cancer that are associated with copy number variations (CNV) or oncogenic transcription factors. Notably, a long intergenic noncoding RNA (lncRNA) GAPLINC (gastric adenocarcinoma predictive long intergenic noncoding RNA) displayed considerable predictive effects, when applied alone or combined with others, in the diagnosis and prognosis of gastric cancer. GAPLINC expression strongly correlated with CD44 in gastric cancer tissues, and the promalignant functions of GAPLINC could be neutralized by suppression of CD44. We provide both in vitro and in vivo data to demonstrate that GAPLINC forms a molecular decoy for miR211-3p, which targets CD44 for degradation. By these efforts, we aim to propose a model for GAPLINC-mediated cell migration and proliferation in gastric cancer.

Materials and Methods

Clinical and histologic evaluation of human tissues
The human specimens in this study were sanctioned by the local ethics committee at the Shanghai Jiao-Tong University School of Medicine Renji Hospital (Shanghai, China). None of the patients received preoperative treatment, including chemotherapy or radiotherapy. The paratumorous tissues were taken at a distance of 2 to 3 cm from the tumor and non-tumorous samples were taken at a distance of at least 5 cm from the tumor, and all tissues were examined histologically. The biopsies of chronic gastritis were obtained from outpatients during endoscopic procedure.

Cell culture
The human MGC803, SGC7901, HCT116, and H1299 cells were obtained from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China) where they were characterized by Mycoplasma detection, DNA-fingerprinting, isozyme detection, and cell vitality detection. These cell lines were immediately expanded and frozen such that they could be restarted every 3 to 4 months from a frozen vial of the same batch of cells. Cells were cultured at 37°C in an atmosphere of 5% CO₂ in RPMI-1640 medium (Invitrogen) supplemented with 10% fetal bovine serum, penicillin, and streptomycin (Thermo Scientific) in 25-ML culture flasks.

Microarray study on lncRNA expression in cancer tissues
Briefly, samples (10 gastric cancer tissues and 10 corresponding nontumor tissues) were used to synthesize double-stranded complementary DNA (cDNA), which was then labeled and hybridized to the 8 × 60 K LncRNA Expression Microarray (ArrayStar). The lncRNA expression microarray used in this study mainly classifies its probes as the following subtypes: (i) enhancer lncRNAs: contain profiling data of all lncRNAs with enhancer-like function (19). (ii) Rinn lincRNAs: contain profiling data of all lincRNAs based on John Rinn's articles (20, 21). (iii) HOX cluster: contains profiling data of all probes in the four HOX loci, targeting 407 discrete transcribed regions, lncRNAs, and coding transcripts (22). (iv) LincRNAs nearby coding gene: contain the differentially expressed lincRNAs and nearby coding gene pairs (distance <300 kb). (v) Enhancer lncRNAs nearby coding gene: contain the differentially expressed enhancer-like lncRNAs and their nearby coding genes (distance <300 kb). After having washed the slides, the arrays were scanned by the Agilent Scanner G2505C. Agilent Feature Extraction software (version 11.0.1.1) was used to analyze acquired array images. Quantile normalization and subsequent data processing were performed using the GeneSpring GX v11.5.1 software package (Agilent Technologies). Data are available via Gene Expression Omnibus (GEO) GSE50710 (ArrayStar LncRNA array; 20 samples).

Microarray for detection of GAPLINC-associated signaling
Total RNA from the human gastric cancer MGC803 cells with GAPLINC stably knockdown and control MGC803 cells were isolated and quantified. The RNA integrity was assessed by standard denaturing agarose gel electrophoresis. The expression profiles were determined using Affymetrix Human Genome U133Plus 2.0 arrays. Quantile normalization and subsequent data processing were performed using the Affymetrix Microarray Suite 5.0 statistical algorithm. Data are available via GEO GSE51651 (Affymetrix HGU133-P2 array; 6 samples).

Sample classification model based on microarray data
We used a widely applied approach to predict gastric cancer from gene expression profiling, based on an enhancement of the simple nearest prototype (centroid) classifier (26). The PAM algorithm shrinks the prototypes, and hence obtains a classifier that is often more accurate than competing methods. The method of “nearest shrunk centroids” identifies subsets of genes that best characterize each class. The shrinkage consists of moving the centroid toward zero by a threshold, which is determined according to the prediction error of the model. As the threshold increases, the number of genes left in the model decreases. To guide the choice of threshold, PAM does K-fold cross-validation for a range of threshold values. It chooses the highest threshold (i.e., the least genes), given the same prediction error.

(Additional Materials and Methods in Supplementary Data).

Results

Deregulation of lncRNAs is associated with recurrent CNVs in gastric cancer
To obtain the transcriptional profiles for both lncRNAs and mRNAs in gastric cancer, paired gastric cancer tissues and normal tissues (n = 20) were analyzed using ArrayStar lncRNA microarray. When the criteria P < 0.05 and fold change >1.5 was adopted, we found similar numbers of lncRNAs being significantly upregulated (n = 659) or downregulated (n = 709) in gastric cancer (plotted in Fig. 1A; detailed gene information shown in Supplementary Table S1). As the alteration of gene expression is often associated with CNV in cancers, we tested whether CNV is also prevalent in deregulated lncRNAs in gastric cancer. To this end, the genomic regions with CNV in gastric cancers were obtained from the Cancer Genome Atlas (TCGA) dataset, and then re-assigned to lncRNA gene loci using the CNTools algorithm. In all 659 upregulated lncRNAs, 215 (32.6%) were mapped to genomic loci with gained CNVs in gastric cancers (Fig. 1A; data in Supplementary Table S2),
Figure 1. Deregulated lncRNAs predict gastric cancer from normal tissues. A, an overview of deregulated lncRNAs in gastric cancer mapped with recurrent CNVs and cancer-related transcription factors. The histogram in blue shows all deregulated lncRNAs in gastric cancer (criteria: $P < 0.05$ and fold change $>1.5$; gene names labeled in outermost layer). The scatter plot shows CNVs of genes encoding lncRNAs (outer layer for amplification and inner layer for deletion). The upregulated lncRNAs that associate with gained CNV are labeled in red above the histogram. The links in the innermost layer indicate preferential binding of two transcription factors (mutant p53 and STAT1) to the promoters of upregulated lncRNAs in gastric cancer. B, misclassification error curves of predictive models using mRNAs (top) and lncRNAs (bottom) in the cross-validation process. The error rates for classification of cancerous and normal tissues are plotted in red and green, respectively. The bottom x-axis shows the threshold for shrinking the centroids (parameter for classification algorithm), and the top x-axis indicates the number of genes left in the model (corresponding to each threshold value). Both prediction models reached the lowest error rates of 0.196. C, differential expression of lncRNAs included in the predictive model for gastric cancer. Genes were ranked by the differences in their average expression levels in normal (green) and cancer (red) tissues. GAPLINC was recognized as the most upregulated lncRNA included in the predictive model. D, scatter plot showing the expression levels of predictive lncRNAs in normal (circles in green) and cancer (red) samples. GAPLINC was recognized as the most upregulated lncRNA included in the predictive model. E, relative expression level of GAPLINC using real-time PCR in 48 paired normal gastric tissues and gastric cancer tissues, which indicated significantly higher expression level of GAPLINC in gastric cancer tissues by paired t test ($P < 0.0001$). F, ROC curve for prediction of gastric cancer using RT-qPCR-based GAPLINC expression level. The AUC was 0.714, with 95% CI and $P$ value indicated.
suggesting that gained CNV may contribute to upregulation of lncRNAs in gastric cancer.

**Preferential binding of transcription factors to upregulated lncRNAs**

Because transcriptional factors play central roles in controlling the initiation of gene expression, thus we tried to search for transcription factors that might be linked to lncRNA deregulation. Recent advances in chromatin immunoprecipitation sequencing (ChIP-seq) provide unbiased and comprehensive knowledge of transcription factor binding patterns throughout the genome, and the growing archive of ChIP-seq data have included many important DNA-binding proteins. To search for transcription factors that may contribute to the upregulation pattern of lncRNAs in gastric cancer, we set to analyze preferential binding of transcription factors to the promoters of upregulated lncRNAs. To this end, we obtained ChIP-seq data of 97 transcription factors from the ENCODE project and GEO database, and analyzed their binding to promoters of differentially expressed lncRNAs (criteria: <5 kb upstream of transcription starting site). Interestingly, a few transcription factors (namely mutant p53, STAT1, and BCL3) seemed to bind preferential to the promoters of upregulated lncRNAs (Fig. 1A and Supplementary Fig. S1A; detailed binding lncRNAs listed in Supplementary Table S3). According to the mRNA levels revealed by microarray, these transcription factors were also upregulated in gastric cancer (Supplementary Fig. S1B). Experimental validation using qPCR suggested that mutant p53 and STAT1 could indeed upregulate the bound lncRNAs, while BCL3 displayed more varied effects. We further tested the effect of mutant p53 on one of the bound lncRNA uc002kmd.1, and found that ectopic expression of mutant p53 R248W could indeed promote the expression of uc002kmd.1 in MGC803 gastric cancer cells, HCT116 colorectal cancer, and H1299 lung cancer (p53-null) cells (Supplementary Fig. S2A). In addition, gastric cancer tissues carrying missense p53 mutations also expressed higher levels of uc002kmd.1 (Supplementary Fig. S2B). The uc002kmd.1 promoter contains mutant p53-binding motifs that have been revealed previously (27), and ChIP assay confirmed that mutant p53 R248W could bind to uc002kmd.1 promoter in vivo (Supplementary Fig. S2C and S2D). In fact, mutation of p53 is one of the most common steps in gastric carcinogenesis (28), and the involvement of STAT1 proto-oncogene in gastric cancer has also been suggested by multiple studies (29, 30). These findings suggest that cancer-related transcription factors may participate in modulating the expression patterns of lncRNAs in gastric cancer, analyzing ChIP-seq data seems to be useful for interpreting the potential effects of transcription factors in lncRNA expression.

**LncRNA-based gastric cancer sample prediction**

The PCG expression profiles have been thoroughly investigated for their abilities to diagnose cancers or discriminate between cancer types, but the efficacy of lncRNAs for such purposes has rarely been reported. Here, we applied a widely used "nearest shrunken centroid method" to classify gastric cancer and normal tissues according to their lncRNA or mRNAs expression profiles. Interestingly, lncRNAs displayed equal predictive power as mRNAs on discriminating carcinous and normal tissue (lowest error rate = 0.196 for both sets; Fig. 1B). The trained prediction signature included nine lncRNAs, wherein the most upregulated was uc002kmd.1 (Entrez gene ID: AX721193; Fig. 1C and D), a lincRNA sitting on the shorter arm of chromosome 18 (924-bp long). We used real-time quantitative PCR (RT-qPCR) to quantify the level of uc001kmd.1 in 48 normal gastric mucosa and paired gastric cancer mucosa, and confirmed the significant upregulation of uc002kmd.1 in gastric cancer ($P$ < 0.0001, Fig. 1E). Furthermore, receiver operating characteristic (ROC) curves were determined to evaluate the sensitivity and specificity of uc002kmd.1 expression in predicting gastric cancer tissues from normal tissues. Notably, uc002kmd.1 displayed considerable predictive significance, with an area under curve (AUC) of 0.714 (Fig. 1F). Given the cancer-predictive value of this RNA, it is hereafter referred to as GAPLINC.

**GAPLINC upregulation associates with shorter survival of gastric cancer patients**

To test whether GAPLINC expression is correlated with poor prognosis of gastric cancer, the expression level of GAPLINC was evaluated by in situ hybridization (ISH) in 90 patients with gastric cancer with different clinicopathologic features (Fig. 2A and B). GAPLINC level was higher in gastric cancer tissues compared with paired normal gastric tissue based on ISH (Fig. 2C). The patients with gastric cancer were then stratified according to GAPLINC expression level (median split) and compared for different clinicopathologic features (age, sex, tumor size, lymph node status, distant metastasis, and survival time). The average tumor size in the GAPLINC-high expression group was significantly larger than that in the GAPLINC-low expression group (Fig. 2D; Mann–Whitney test, $P$ = 0.0097). Moreover, the occurrence of severe lymph node invasion was more frequent in the GAPLINC-high expression group ($\chi^2$ test, $P$ = 0.0319; Fig. 2E). In addition, high expression of GAPLINC associated with shorter patient survival (Fig. 2F; $P$ < 0.01, Mantel–Cox test), and the association was stronger than a protein marker that we reported previously (synbindin; $P$ = 0.0468, Mantel–Cox test; ref. 3). ROC curves were determined to evaluate the sensitivity and specificity of the survival prediction based on the lncRNA ISH intensity and the American Joint Committee on Cancer (AJCC) stages (Fig. 2G). Interestingly, the AUC for GAPLINC-based prediction was higher than AJCC-based prediction (0.758 vs. 0.682), and combined both indexes could further improve the survival prediction (AUC, 0.794). The AJCC tumor–node–metastasis (TNM) staging system has been widely accepted as a powerful predictor of treatment response and survival in gastric cancer, thus it is of interest to test whether the prognostic value of the GAPLINC is independent of AJCC stage. Multivariable Cox regression analysis adjusting AJCC stage and other factors confirmed the association between GAPLINC expression and shorter survival [hazard ratio (HR), 1.539; 95% confidence interval (CI), 1.219–1.944; $P$ < 0.01; Supplementary Table S4].
GAPLINC is required for efficient proliferation and invasion of gastric cancer cells

To test whether GAPLINC is required for maintenance of malignant phenotypes of gastric cancer cells, specific siRNAs were used to knockdown GAPLINC expression in two gastric cancer cell strains MGC803 and SGC901, which express higher level of GAPLINC (Supplementary Fig. S2E). Transwell assay revealed a substantial decrease in the number of cells that penetrated the porous filter, suggesting impaired invasion ability for both cell lines (Fig. 3A–C). Meanwhile, cDNA-mediated ectopic expression of GAPLINC significantly increased the invasiveness of both cell lines.

Figure 2. GAPLINC expression correlates with poor outcome of gastric cancers. A, ISH of GAPLINC in normal gastric mucosa or gastric cancer tissues. Paraffin-embedded tissue sections were stained using specific probe for GAPLINC. B, ISH of U6 spliceosomal RNA in gastric cancer tissues or normal gastric mucosa as control. Paraffin-embedded tissue sections were stained using specific probe for U6 in purple-blue. C, statistical analysis of GAPLINC expression in 90 paired normal and cancerous gastric tissues. The y-axis indicates staining intensity of GAPLINC. The expression level of GAPLINC was significantly higher in cancerous tissues (P < 0.0001, paired t test). D, statistical analysis of the size of gastric cancers in GAPLINC-low and -high expression groups. The average tumor size in two groups was compared using the Mann–Whitney test (P < 0.0001). E, relevance of GAPLINC expression to clinicopathologic features of gastric cancers. The patients were classified into two groups according to GAPLINC expression levels. The numbers of patients, distant metastasis, average tumor volume, and severe invasion into lymph nodes (3 of 5 affected) are displayed in each group. The P values indicating statistical significance of difference between the two groups are also shown in the table. F, survival of patients in GAPLINC-low expression group and -high expression group. The survival time of patients after surgery was compared between groups using the Mantel–Cox test, which indicated significantly longer survival of patients in the GAPLINC-low expression group (P < 0.0001). G, ROC analysis of ISH-based GAPLINC expression level for survival prediction of patients with gastric cancer.
The CCK8-based viability assay detected significant decrease in the proliferation of MGC803 and SGC901 cells after knockdown of GAPLINC, while overexpression of GAPLINC dramatically promoted the cells proliferation (Fig. 3D and E and Supplementary Fig. S3). Flow cytometry assay indicated that suppression of GALINC induced the increase of cell apoptosis by phycoerythrin (PE)-conjugated Annexin V staining and FACS (Fig. 3E–H). These findings suggest that GAPLINC may not only be a potential marker, but also play a driving role in gastric cancer development.

GAPLINC regulates cell invasion by controlling CD44 expression

To probe the GAPLINC-associated pathway on an unbiased basis, we investigated the gene expression profiles of gastric cancer cells that were suppressed for GAPLINC expression (schematic shown in Fig. 4A). To this end, the MGC803 cells were treated with specific siRNAs for GAPLINC, and the levels of all mRNAs were measured by Affymetrix Human Genome U133 Plus 2 microarrays (triple repeats for each condition; data accessible via GEO #GSE51651). The GAPLINC-associated pathways were determined by gene set enrichment analysis.
Figure 4. GAPLINC promotes cell migration by regulating CD44. A, a schematic flowchart showing the process of microarray study on GAPLINC-associated pathways. Gastric cancer cells were treated by expression vectors encoding siRNAs for GAPLINC or control siRNAs, and the mRNA expression profiles were determined by microarray. The combination of GSEA and gene expression correlation study identified CD44 in the cell migration pathway as potential regulatory target of GAPLINC. B, knockdown of GAPLINC caused alteration in multiple pathways, and “regulation of cell migration” pathway was found with the highest significance. This is based on the following principle: when most genes in a defined pathway (gene set) are affected, a higher enrichment score is assigned to that pathway. C, enrichment plot of the cell migration pathway in the GSEA analysis. All genes were ranked by their changes in association with GAPLINC knockdown (filled gray curve in the bottom), and the positions of migration-associated genes in the list are labeled with vertical lines (middle). Most genes in the migration pathway belonged to the upregulated set (heatmap red and blue). The P value for GSEA analysis is provided in the enrichment plot (top). D, correlation between the expression of GAPLINC and CD44 revealed by microarray study (P < 0.0001, Pearson correlation). E, the levels of GAPLINC and CD44 were measured by RT-qPCR, and their correlation was determined by Pearson correlation analysis (P < 0.0001; R = 0.827). F and G, modulating GAPLINC expression significantly affected CD44 mRNA expression. Gastric cancer cells were stably transfected with either GAPLINC shRNA/control shRNA (F), or GAPLINC cDNA/control vector (G), followed by detection of CD44 mRNA using RT-qPCR. Knockdown of GAPLINC decreased CD44 expression (MGC803, P < 0.01; SGC7901, P < 0.001, Student t test), while ectopic expression of GAPLINC increased CD44 mRNA level (MGC803, P < 0.01; SGC7901, P < 0.001, Student t test). H, the expression of GAPLINC was suppressed or enhanced as described above, and the expression level of CD44 protein was determined by Western blot analysis. The level of α-tubulin was also detected as loading control. I and J, the proinvasion effect of GAPLINC is neutralized by suppressing CD44 expression. Gastric cancer cells were transfected with GAPLINC cDNA in the presence or absence of siRNAs for CD44, and the invasion ability of cells was determined by Transwell assay. Statistical result based on three independent experiments is indicated in I, and representative Transwell cell staining images are shown in J. The magnified sections are shown in the bottom.
(GSEA), which determines whether different pathways (sets of genes) show statistically significant differences between two biologic states (31). Among the significantly affected pathways, "regulation of cell migration" was assigned with the highest enrichment score (Fig. 4B and C). By analyzing the correlation between GAPLINC and miRNAs in the IncRNA microarray dataset (containing 10 tumors and 10 normal tissues), we found CD44 with the highest correlation coefficient in this pathway (Pearson correlation, $R = 0.810; P < 0.0001$; Fig. 4D). We used reverse transcription quantitative PCR (RT-qPCR) to measure the expression of GAPLINC and CD44 in 15 gastric cancer tissues, 15 normal tissues, and 27 chronic gastritis tissues (Fig. 4E). The result confirmed the strong correlation between GAPLINC and CD44 (Pearson correlation, $R = 0.827; P < 0.0001$).

Next, we manipulated GAPLINC expression in gastric cancer cells and monitored its effect on CD44 expression. Knockdown of GAPLINC by specific siRNAs substantially decreased CD44 expression, whereas ectopic expression of GAPLINC caused elevation in CD44 mRNA level (Fig. 4F and G). Consistently, Western blot analysis revealed that knockdown/overexpression of GAPLINC respectively decreased/increased CD44 protein level in gastric cancer cells (Fig. 4H). By Transwell assay, we found that the proinvasion behavior of GAPLINC is neutralized by knockdown of CD44 (Fig. 4l and J). Consistently, the suppression effect of GAPLINC knockdown on cell invasion could be rescued by ectopic expression of CD44 (Supplementary Fig. S4). These findings suggest that GAPLINC confers CD44-dependent effects in gastric cancer cells.

**GAPLINC regulates CD44 expression by competing for miR211-3p**

To probe the mechanism for GAPLINC-regulated CD44 expression, we firstly tested whether GAPLINC could modulate the transactivation of CD44 mRNA. The promoter of CD44 was cloned upstream of a luciferase reporter gene, and the resultant construct was cotransfected with GAPLINC in gastric cancer cells. As a result, GAPLINC did not affect the transactivation of CD44 promoter (Fig. 5A), suggesting that GAPLINC may regulate CD44 mRNA after it is transcribed.

Because many IncRNAs function as natural "microRNA sponges" that protect mRNAs by competing for their targeting microRNAs, it is worthy to test if GAPLINC may play such a role. Interestingly, a microRNA, namely miR211-3p, was predicted to target both CD44 and GAPLINC, with remarkable binding energy estimated by the widely used RNAup algorithm as in Fig. 5B (32). To validate the effects of miR211-3p, we cloned the 3′-UTR of CD44, mutant 3′-UTR of CD44 and GAPLINC downstream of a luciferase gene, and cotransfected these reporters with miR211-3p mimics in gastric cancer cells. As expected, miR211-3p significantly decreased the luciferase signals of both reporters (Fig. 5C and D). However, miR211-3p had no effect on mutant 3′-UTR of CD44 (Fig. 5E), suggesting the specificity of miR211-3p binding site on CD44 3′-UTR. Furthermore, treatment of gastric cancer cells by miR211-3p mimics significantly decreased CD44 and GAPLINC RNA levels (Fig. 6A and B), confirming that miR211-3p could target GAPLINC and CD44. The effect of miR211-3p on CD44 was also confirmed by Western blot analysis (Fig. 6C).

Furthermore, we cloned 3′-UTR of CD44 into a luciferase reporter and cotransfected the construct with GAPLINC siRNA or control siRNA. As a result, knockdown of GAPLINC significantly reduced the luciferase intensity, suggesting that GAPLINC is required for the abundant expression of CD44 (Fig. 6D). Interestingly, the pro-CD44 effect of GAPLINC could be neutralized by miR211-3p in MGC803 and SGC7901 cells (Fig. 6E–G), suggesting that GAPLINC confers miR211-3p-dependent effects on CD44 expression. Finally, we studied the relationships between miR211-3p and its target RNAs (both CD44 and GAPLINC) in tumor tissue samples. The expression level of miR211-3p negatively correlated with CD44 and GAPLINC RNA levels (Pearson, $R = -0.33$ and $-0.25$, respectively), which strongly supported our notion that miR211-3p negatively regulates CD44 and GAPLINC in gastric cancers (Fig. 6H and I).

**GAPLINC correlates with CD44 activation in gastric cancer tissues**

To confirm the impact of GAPLINC on CD44 in vivo, we generated xenograft models by implanting MGC803\textsuperscript{GAPLINC-KD} (stably knockdown GAPLINC) or the control MGC803\textsuperscript{vector} cells into nude mice. As a result, suppression of GAPLINC produced a marked decrease in the rate of xenograft subcutaneous tumor growth (Fig. 7A–C, Supplementary Table S5). Intriguingly, knockdown of GAPLINC dramatically decreased the level of CD44 (Fig. 7D and E) and increase the level of miR211-3p in xenograft tumors (Fig. 7F). These data confirmed the effect of GAPLINC on CD44 activation and strongly support our notion that GAPLINC contributes to the malignant phenotype by activating CD44.

**Discussion**

Despite recent progresses in discovering cancer-related IncRNAs, few IncRNAs have been characterized for their exact roles in gastric cancer. Our study has provided a landscape of IncRNA deregulation mapped with associated CNVs and transcription factors, and this may facilitate further exploration of functional IncRNAs in gastric cancer. The genes with causal roles in gastric cancer are often located in the genomic regions CNV recurrence, and transcription factors binding might be beneficial for discovering cancer-related IncRNAs.
We have presented a proof-of-principle study for the prediction of cancer/normal tissues using lncRNA expression profiles. Previous efforts have thoroughly explored the feasibility of using mRNA expression profiles for predicting cancer or discriminating between cancer subtypes, but lncRNA-based predictive biomarkers are rarely reported. This might be due to the assumption that mRNAs are more likely to be functional, and thus may reflect the status of various biologic pathways. However, our results suggest that lncRNA expression signatures are at least comparable with mRNAs in sample stratification. In fact, lncRNAs have an obvious merit of their relatively simple functional layout as transcriptional levels.

In many cancers, the functions of PCGs may also be affected by mutations and protein translation/modifications. Taking the TP53 gene as an example, the upregulation of p53 mRNA or protein in cancer often associates with somatic mutation of TP53 (occurring in 50% cancers), rather than functional activation of this tumor-suppressive pathway (35). In the form of measured RNA expression (by PCR or hybridization-based assays), lncRNAs may better reflect the biologic status of cancer cells than PCGs, and thus should be further studied as predictive biomarkers.

Importantly, GAPLINC was the most upregulated gene in the predictive model containing nine lncRNAs, and its expression alone could also predict gastric cancer with considerable accuracy. GAPLINC expression is required for the efficient invasion and proliferation of gastric cancer cells, suggesting the functional involvement of GAPLINC in gastric cancer. By suppression of GAPLINC and microarray study, we identified a robust correlation between GAPLINC and CD44, a well-known oncogene in gastric cancer.

![Figure 5. Dual roles of miR211-3p in regulating GAPLINC and CD44.](image-url)

A. GAPLINC displays no effect on the transactivity of CD44 promoter. The promoter of CD44 was cloned upstream of a luciferase reporter gene, and the resultant construct was cotransfected with GAPLINC in gastric cancer cells. B. The RNAup algorithm predicted potential binding of miR211-3p to GAPLINC and to CD44, with considerable sequence complementary in the indicated regions (binding energy indicated below). C and D, both GAPLINC and CD44 are targeted by miR211-3p. The 3' UTR of CD44 mRNA and GAPLINC were respectively inserted downstream of a luciferase gene. The reporter vector was cotransfected with a Renilla luciferase vector (for normalization) to gastric cancer cells, which were treated by miR211-3p mimics or control mimics. The luciferase signals (firefly/Renilla) of both CD44 3' UTR (C) and GAPLINC reporter (D) genes were significantly decreased when cells were treated with miR211-3p (P values indicated, Student t test). E, mutant 3' UTR of CD44 is not affected by miR211-3p. The 3' UTR of CD44 was mutated on the predicted binding site that is shown in B and was tested in the luciferase assay as described above. The result shows that miR211-3p did not alter the luciferase signal of the mutant CD44 3' UTR (P values indicated, Student t test).
characterized gene involved in cancer proliferation, migration, and angiogenesis. Both \textit{in vitro} and \textit{in vivo} data demonstrated that GAPLINC regulates CD44 expression by competing for miR211-3p, which targets CD44 for degradation (model illustrated in Fig. 7G). We have also demonstrated by xenograft models that targeting of GAPLINC could suppress tumor growth.

Figure 6. GAPLINC regulates CD44 via competing for miR211-3p. A and B, MGC803 (A) and SGC7901 (B) cells were treated with miR211-3p or control mimics, followed by detection of GAPLINC and CD44 levels using RT-qPCR. Treatment by miR211-3p significantly decreased the levels of CD44 mRNA and GAPLINC in both MGC803 (A) and SGC7901 cells (B). C, the expression of miR211-3p was enhanced as described above, and the expression level of CD44 protein was determined by Western blot analysis. The level of \(\alpha\)-tubulin was also detected as loading control. D, GAPLINC is required for the stability of CD44 3'-UTR. The 3'-UTR sequence of CD44 mRNA was inserted downstream of a luciferase gene. The reporter vector was cotransfected with a \textit{Renilla} luciferase vector (for normalization) to gastric cancer cells, which were treated with GAPLINC siRNA or control siRNA. The luciferase signal (\textit{firefly}/\textit{Renilla}) of reporter gene was significantly decreased when cells were treated with GAPLINC siRNA (\(P\) values indicated, Student \(t\) test). E and F, gastric cancer cells were transfected with GAPLINC cDNA in the presence or absence of miR211-3p mimics, and level of cDNA was determined by RT-qPCR in both MGC803 (E) and SGC7901 (F) cells. G, cells were treated as described above, and CD44 protein level was determined by Western blot analysis. The \(\alpha\)-tubulin protein was detected as loading control. H and I, the correlation between miR211-3p and CD44/GAPLINC in 15 gastric cancer tissues, 15 normal tissues, and 27 chronic gastritis tissues. Expression levels of miR211-3p, CD44, and GAPLINC were determined by RT-qPCR, and Pearson correlation was used to analyze the relationships between miR211-3p and CD44 (H), and between miR211-3p and GAPLINC (I). \(P\) values are indicated in the plot.
Figure 7. Targeting GAPLINC decreased CD44 expression and tumor growth in vivo. A and B, knockdown of GAPLINC inhibited growth of xenograft in nude mice. Human gastric cancer MGC803 cells with GAPLINC stable knockdown or negative control were injected subcutaneously into the left hip to establish xenograft model. At the last time point (17 days after first injection), tumors in both groups were measured both in situ (A) and after resection (B), shown for both groups. C, statistical analysis of tumor volume in MGC803GAPLINC-KD (knockdown) and MGC803vector (control) groups (values also shown in Supplementary Table S5). Knockdown of GAPLINC significantly delayed tumor progression as compared with the control group. D, RT-qPCR of GAPLINC and CD44 in xenografts treated with GAPLINC stable knockdown clone and control clone. It can be observed that knockdown of GAPLINC dramatically decreased the level of CD44. E, Western blot analysis of xenografts in MGC803GAPLINC-KD and MGC803vector groups suggest decreased expression of CD44 protein upon stable knockdown of GAPLINC. F, RT-qPCR of miR211-3p in xenografts in MGC803GAPLINC-KD and MGC803vector groups. It can be observed that knockdown of GAPLINC dramatically increased the level of miR211-3p (P value indicated, Student t test). G, schematic representation showing GAPLINC-mediated cell proliferation and invasion in gastric cancer. The miR211-3p has dual target specificity for both GAPLINC and CD44, which leads to degradation of the bound RNA. The high expression of GAPLINC competes to bind miR211-3p, which reduces the degradation of CD44 mRNA and leads to increased translation of the functional CD44 protein. The CD44 oncoprotein has well-established roles in promoting cancer proliferation, migration, and angiogenesis, which mediates the oncogenic effects of GAPLINC in gastric cancer.
progression in vivo. These findings reveal a regulatory network involving different forms of RNAs that is imbalanced in gastric cancers, and suggest GAPLINC as a potential biomarker and therapeutic target for this deadly disease.

In conclusion, our study identifies GAPLINC as a potential biomarker in gastric cancer and reveals its pivotal regulatory effect on the expression of CD44 oncoprotein.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Disclaimer

The sponsors of this study had no role in the collection of the data, the analysis and interpretation of the data, the decision to submit the manuscript for publication, or the writing of the article.

Authors’ Contributions

Conception and design: Y. Hu, J. Tang, Y. Chen, J. Xu, J.-Y. Fang

Development of methodology: Y. Hu, J. Xu

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): Y. Hu, J. Wang, J. Qian, X. Kong, Y. Wang

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): Y. Hu, J. Wang, H. Chen, J. Hong, W. Zou, J. Xu

Writing, review, and/or revision of the manuscript: Y. Hu, J. Xu, J.-Y. Fang

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): Y. Hu

Study supervision: J. Xu, J.-Y. Fang

Acknowledgments

The authors thank Wenxin Qin and Nan Shen for their helpful comments on this article.

Grant Support

This project was supported by grants from the National Natural Science Foundation of China (30971330, 31371428, 8120180024, 81000861, 81322036, 81302230, 81372230, and 81372238), the Program for Innovative Research Team of Shanghai Municipal Education Commission, the Shanghai ‘Oriental Scholars’ project (2013XJ), the Shanghai Science and Technology Commission ‘Pujiang Project’ (13PJ105090), and the Shanghai Natural Science Foundation (12ZR1417900).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 10, 2014; revised August 19, 2014; accepted September 5, 2014; published OnlineFirst October 2, 2014.

References


Long Noncoding RNA GAPLINC Regulates CD44-Dependent Cell Invasiveness and Associates with Poor Prognosis of Gastric Cancer

Ye Hu, Jilin Wang, Jin Qian, et al.

Cancer Res  Published OnlineFirst October 2, 2014.

Updated version  Access the most recent version of this article at: doi:10.1158/0008-5472.CAN-14-0686

Supplementary Material  Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2014/10/03/0008-5472.CAN-14-0686.DC1

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/early/2014/11/19/0008-5472.CAN-14-0686. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.