Genome-Wide Identification and Characterization of DNA Methylation and Long Non-Coding RNA Expression in Gastric Cancer

Peng Song[1†], Lei Wu[2†] and Wenxian Guan[1*]

[1] Department of General Surgery, Nanjing Drum Tower Hospital, The Affiliated Hospital of Nanjing University Medical School, Nanjing, China, [2] Department of Laboratory Medicine, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

Abnormal DNA methylation, an epigenetic modification, has increasingly been linked to the pathogenesis of many human cancers. However, there has been little focus on the DNA methylation patterns of genes encoding long noncoding RNAs (lncRNAs) in gastric cancer (GC). This study comprehensively determined DNA methylation and lncRNA expression profiles in GC through genome-wide analysis. Differentially methylated loci and lncRNAs were identified by integrating multi-omics data. In total, 548 differentially methylated CpG sites in lncRNA promoters and 2,399 differentially expressed lncRNAs were screened that were capable of distinguishing GC from normal tissues. Among them, 22 differentially methylation sites in 17 lncRNAs were inversely related to expression levels. Further analysis of DNA methylation status and gene expression level in GC revealed that three CpG sites (cg01550148, cg22497867, and cg20001829) and two lncRNAs (RP11-366F6.2 and RP5-881L22.5) were significantly associated with GC patient overall survival. Molecular function analysis showed that these abnormally methylated lncRNAs were mainly involved in transcriptional activator activity. Our study identified several lncRNAs regulated by aberrant DNA methylation that have clinical utility as novel prognostic biomarkers in GC. These findings help improve the understanding of methylated patterns of lncRNAs and further our knowledge of the role of epigenetics in cancer development.

Keywords: DNA methylation, long non-coding RNA, epigenetics, prognosis, gastric cancer

INTRODUCTION

Gastric carcinoma (GC) is the fourth most prevalent malignancy and third leading cause of cancer death worldwide (Torre et al., 2015). Histologically, GC demonstrates marked heterogeneity at the cytologic level, resulting in the classification of tumor subtypes. Distinct molecular genetic profiles, morphology, and expression of specific markers have been used to investigate the diversity and characteristics of GC (Zouridis et al., 2012; Cancer Genome Atlas Research Network, 2014). Therefore, identifying potential biomarkers to further understand the pathogenesis of GC is critical.

Long noncoding RNAs (lncRNAs) are loosely defined as RNAs more than 200 bases in length with no apparent coding capacity (Mattick and Rinn, 2015). LncRNAs regulate gene expression at transcriptional and post-transcriptional levels and thus are involved in diverse biological functions. Furthermore, recent studies have demonstrated a role for lncRNAs in carcinogenesis (Spizzo et al., 2012; Zhuo et al., 2019). DNA methylation, a key epigenetic mechanism, plays a crucial role in the regulation of gene expression, genomic imprinting, genome stabilization, and chromatin modification. Aberrant DNA methylation has been reported to be involved in the formation and progressions of diseases, especially cancers (Guo et al., 2018; Xu et al., 2019). Recent studies have showed that expression alterations of lncRNA-encoding genes mediated by changes in methylation can subsequently affect their downstream targets. For instance, the lncRNA C5orf66-AS1 functions as a tumor suppressor gene in GC, and aberrant hypermethylation of the regions around its transcription start site (TSS) is associated with its expression and is cancer-specific (Guo et al., 2018). This study indicated that hypermethylation of the C5orf66-AS1 promoter may serve as a potential prognostic marker in predicting GC patient survival. Shahabi et al. identified an epigenetically deregulated lncRNA linc00261, whose expression was lost in lung adenocarcinoma through DNA methylation silencing. The authors found that linc00261 acted upstream of ATM activation to facilitate DNA damage response activation and its loss resulted in malignant phenotypes and predisposed lung cells to cancer development (Shahabi et al., 2019). Additionally, lncRNAs showing aberrant DNA methylation may serve as potential epigenetically-based diagnostic factors. Silencing from CpG-island methylation of promoter-induced transcribed ultraconserved regions (T-UCRs) is common in many tumors and linked to colorectal cancer diagnosis (Kottorou et al., 2016; Honma et al., 2017). Therefore, elucidating the relationship between DNA methylation and lncRNA expression is essential for understanding GC development and potentially identifying new prognostic or diagnostic markers.

Here, we employed multigenomic data from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) datasets to systematically characterize global DNA methylation levels, lncRNA expression profiles, and clinical features in GC. Our results decode the landscape of DNA methylation-mediated regulation for lncRNAs and provide promising biomarkers in the diagnosis and treatment of GC.

METHODS

DNA Methylation and Gene Expression Data

The DNA methylation array data (Illumina Infinium Human Methylation27, 450 BeadChip) were downloaded from the UCSC Xena browser (https://xenabrowser.net/). A Human Methylation27 BeadChip array of GC (GSE30601) was obtained from the GEO database (https://www.ncbi.nlm.nih.

gov/geo/). Level-3 RNA-sequencing data (HTSeq-Counts and HTSeq-FPKM-UQ) and the clinicopathological and survival data of GC patients were also downloaded from the Xena website.

Analysis of DNA Methylation Data

Differentially methylated CpG sites (DMCs) and differentially methylated regions (DMRs) between GC samples and adjacent tissues were identified using the minfi package (Version: 1.32.0; http://www.bioconductor.org/packages/release/bioc/html/minfi.html) (Fortin et al., 2017). Bump hunting method was applied to identify DMRs. False discovery rate (FDR) was calculated from multiple testing corrections of raw P-value by the Benjamini and Hochberg method (Benjamini et al., 2001). The genomic annotation of each CpG site was conducted using the hm27.hg38.manifest file (http://zwdzwd.io/InfiniumAnnotation). The coordinates of the individual lncRNA were extracted from GENCODE v22 (https://www.gencodegenes.org/human/release_22.html). After the preprocessing the coordinates of CpG sites and lncRNAs, we further integrated both information based on the genomic location, considering differentially methylated loci within promoter regions (DNA sequences between −2,500 and 1,000 bp relative to the putative TSS). Manhattan plot was constructed to depict the distribution of CpG sites according to FDR *via* qqman package (Version: 0.1.4; https://cran.r-project.org/web/packages/qqman/index.html) (Turner, 2018).

Differential Long Non-Coding RNA Expression Analysis

Read count tables were imported into the edgeR package for identifying differentially expressed transcripts (Version: 3.7, https://bioconductor.org/packages/release/bioc/html/edgeR.html) (McCarthy et al., 2012). LncRNA catalogue was retrieved from GENCODE v22. Genes with FDR < 0.05 and absolute fold change (FC) > 2 were considered differentially expressed lncRNAs (DElncs).

Integrated Analysis of DNA Methylation and Long Non-Coding RNA Expression

The correlation analysis between DMCs and DElncs was calculated and those with |coefficient of correlation| > 0.3 and P-value < 0.05 were considered significant. The visualization of the potential regulation of CpG sites to genes was constructed in Cytoscape 3.7.1 (Shannon et al., 2003).

Functional Annotation and Enrichment Analysis for Long Non-Coding RNAs

ClusterProfiler tool (Version: 3.8.1, https://bioconductor.org/packages/release/bioc/html/clusterProfiler.html) was used to perform Gene Ontology (GO) function and Gene Set Enrichment Analysis (GSEA) for DElncs with the DMCs (Yu et al., 2012). Spearman's correlation coefficients of expression levels between DElncs and protein-coding RNAs were calculated. The deregulated protein-coding genes were considered for GO analysis, setting parameters as "pAdjustMethod" = "BH," "pvalueCutoff" = 0.05, and "qvalueCutoff" = 0.05 for multiple

comparisons. Terms from GO (molecular function) database slice were tested for enrichment. Mappings between GO terms and Entrez Gene IDs relied on the regularly updated R package org.Hs.eg.db (Version 3.10.0; https://bioconductor.org/packages/release/data/annotation/html/org. Hs.eg.db.html). The value of $\log_2$ (FC) calculated by edgeR package was used as ranking metric for GSEA. We used the canonical pathways sub-collection of the C2 collection in the Molecular Signatures Database as the gene sets in the analysis. The leading-edge subset of genes in an enriched gene set are defined as those that appear in the ranked list before the point at which the running sum reaches its maximum deviation from zero.

Statistical Analysis

Cox proportional hazard regression analyses were carried out to compare clinical features, DNA methylation, DElncs expression, and GC patients' prognosis. Survival curves were compared using Kaplan-Meier Plotter with log-rank test. All statistical analysis was two-sided and P < 0.05 was defined as statistically significant. Statistical analysis was performed using R programming language v.3.5.3.

RESULTS

Characteristics of the DNA Methylation Pattern in Gastric Cancer

Because the TCGA-450k set contained only two normal samples and the number was too small to reach statistical significance with respect to determining the DNA methylation profile of GC, we used the TCGA-27k set to identify DMCs and DMRs. We obtained 6,404 CpG sites with FDR < 0.05 between 48 GC and 25 non-tumor samples and identified 1,078 DMCs with a delta-beta value > 0.2. A total of 103 DMRs were identified based on the following parameters: resamples = 100, cut off = 0.2, and probe

number ≥ 2. The 103 DMRs included 65 hypermethylated regions and 38 hypomethylated regions.

To identify DNA methylation alterations in lncRNA promoter regions, 3,010 CpG sites were examined. The methylation distribution in lncRNAs showed a V-shaped curve around the TSSs, indicating a relative reduction of the methylation density at the TSS (**Figure 1A**). As shown by Manhattan plot (**Figure 1B**), the CpG sites were distributed in all chromosomes, and 698 probes were found using the threshold of FDR < 0.05. We subsequently validated the DNA methylation patterns of the CpG sites of interest in an independent cohort (GSE30601) and found that 548 probes overlapped with the 698 sites reported in TCGA-27k set (**Supplementary Table S1**).

Characteristics of Long Non-Coding RNA Expression in Gastric Cancer

To determine the lncRNA expression profile in GC, RNA-seq data of 375 GC tumors and 32 normal tissues were retrieved from TCGA. Among the 6,820 lncRNAs, we identified 2,399 DElncs, including upregulated 1,830 lncRNAs and 569 downregulated lncRNAs, using the criteria of FDR < 0.05 and absolute FC > 2 (**Figure 2A**, **Supplementary Table S2**). We then analyzed the categories of the 2,399 DElncs, as shown in **Figure 2B**. Long intergenic non-coding RNAs (lincRNAs) accounted for 54.3% of all DElncs, followed by antisense transcripts (30.3%). The remaining non-coding transcript types were sense_intronic transcripts (4.5%), processed_transcripts (3.7%), and sense_overlapping transcripts (1.3%).

Integrated Analysis of Differential Methylation and Long Non-Coding RNA Expression Data

After the profiles of DNA methylation and lncRNA expression were preprocessed, we combined the two omics data for further analysis. By associating the 548 DMCs to 2,399 DElncs, 31
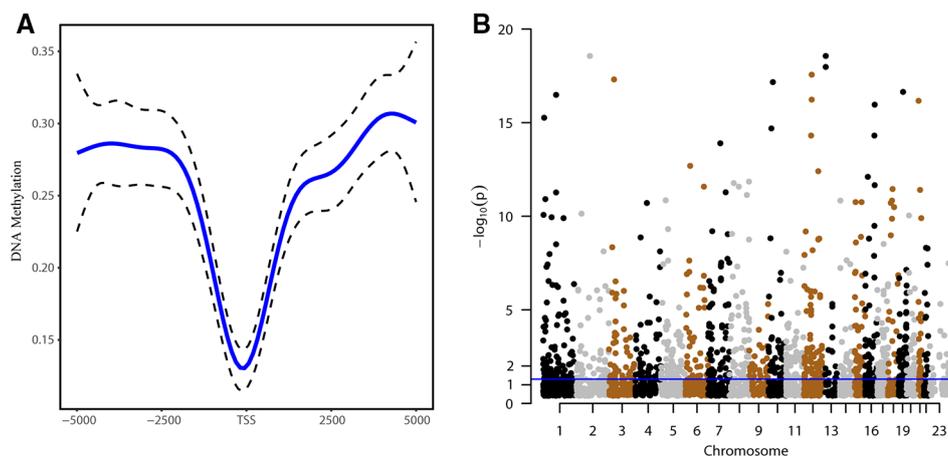

FIGURE 1 | DNA methylation patterns of genes encoding long noncoding RNAs (lncRNAs). **(A)** Distribution of the methylation levels around lncRNA genes in sperm ranging from 5 kb upstream to 5 kb downstream of the transcription start site (TSS). **(B)** Manhattan plot of CpG sites in the promoter regions of lncRNA genes; dots above the blue line indicate CpG sites with P value < 0.05.
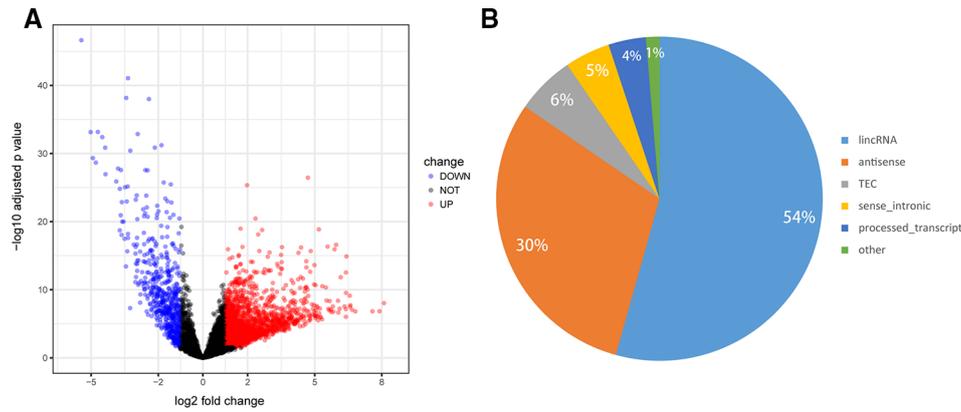
FIGURE 2 | Differential expression profiles of long noncoding RNAs (lncRNAs) in gastric cancer (GC). **(A)** Volcano plot of the differentially expressed lncRNAs between GC tumors and normal tissues. The red points represent lncRNAs that are significantly upregulated in GC while blue points represent downregulated lncRNAs (absolute FC >2 and FDR < 0.05). **(B)** Pie chart shows the number of differentially expressed lncRNAs in each category.

negative correlated pairs and 1 positive correlated pair were obtained in TCGA-27k set. DNA methylation in promoters is well known to negatively correlate with corresponding gene expression (Mosquera Orgueira, 2015). The 31 negative correlated pairs were validated in TCGA-450k set, and 22 probes showed a significantly inverse correlation with the

promoter methylation of 17 aberrantly expressed lncRNAs (**Supplementary Table S3**). We used the negatively correlated pairs to construct a DNA methylation-regulated network that was composed of 39 nodes, including 6 hypermethylated DMCs, 16 hypomethylated DMCs, 13 upregulated lncRNAs, and 4 downregulated lncRNAs (**Figure 3A**). As shown in **Figure 3B**,
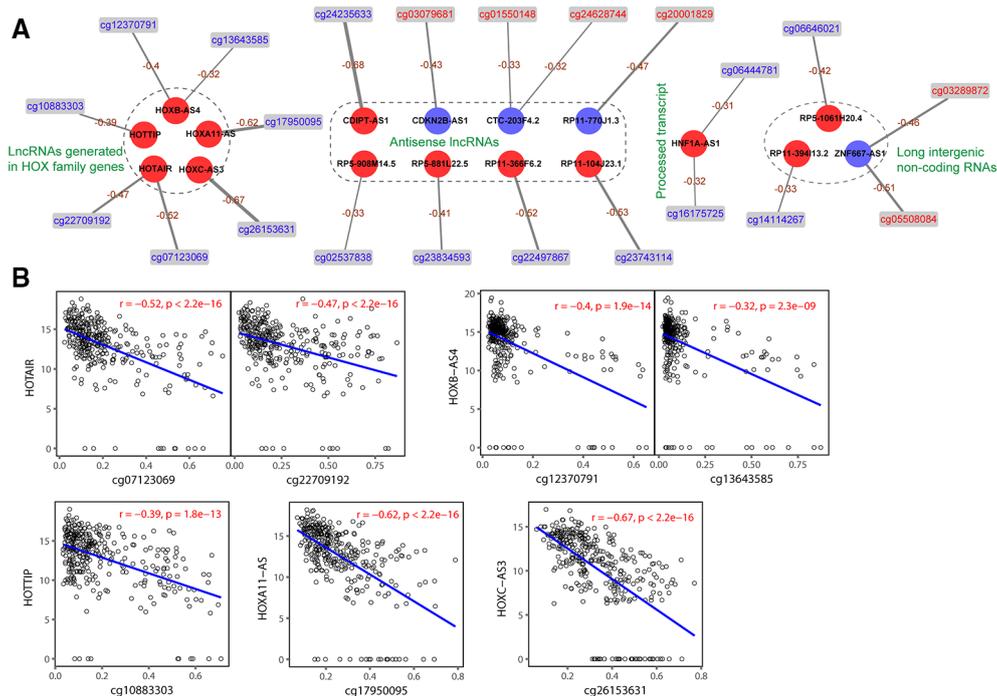

FIGURE 3 | Relation between DNA methylation and long noncoding RNA (lncRNA) expression. **(A)** Correlation between differentially methylated CpGs (DMCs) and lncRNAs. Circles and rectangles represent lncRNAs and DMCs, respectively. Red color indicates upregulated or hypermethylated, and blue indicates downregulated or hypomethylated. **(B)** Correlation (P values derive from Spearman's correlation) between DNA methylation and the expression of HOX family genes associated with five lncRNAs in matched samples.

the expression of 5 lncRNAs (HOTAIR, HOTTIP, HOXA11-AS, HOXB-AS4, and HOXC-AS3) generated in HOX family genes were negatively correlated with their methylation levels. The functions of these 17 lncRNAs are listed in **Supplementary Table S4**; only 6 of the lncRNAs have been reported to function GC (Zhang et al., 2014; Liu et al., 2015; Sun et al., 2016; Wu et al., 2017; Liu et al., 2018; Zhang et al., 2018; Song et al., 2019).

Impact of DNA Methylation and Long Non-Coding RNA Expression on Gastric Cancer Survival

Univariate Cox regression was used to evaluate the association of the 22 probes and 17 lncRNAs with overall survival in GC and the results identified three CpG sites (cg01550148, cg22497867, and cg20001829) and two lncRNAs (RP11-366F6.2 and RP5-881L22.5) with $P < 0.05$. Forest plot demonstrated that the methylation of the three probes and two lncRNAs were associated with the overall survival time of GC patients (**Figure 4A**). **Figure 4B** shows the Kaplan-Meier curves for

survival in GC patients according to RP11-366F6.2 and RP5-881L22.5 expression. High expression of RP11-366F6.2 was significantly correlated with poor survival compared with low expression. Additionally, poor survival was observed for patients with low expression of RP5-881L22.5 compared with patients with high levels ($P = 0.039$). In the multivariate Cox analysis, even after adjustment by tumor stage and other covariates, the expression of the two lncRNAs was still significantly associated with patient survival ($P < 0.05$, **Table 1**).

Association of Deregulated Long Non-Coding RNAs With Biological Pathways and Processes

To better understand the biological function of the 17 DElncs, we constructed a co-expression network of deregulated protein-coding genes and lncRNAs. Using the Spearman's correlation coefficient above 0.8, a total of 32 deregulated mRNAs co-expressed with five lncRNAs (ZNF667-AS1, RP5-881L22.5, HOTAIR, HOTTIP, and HOXC-AS3) were acquired for GO enrichment analysis. Only nine protein-coding genes associated with HOTAIR, HOXC-AC3, and
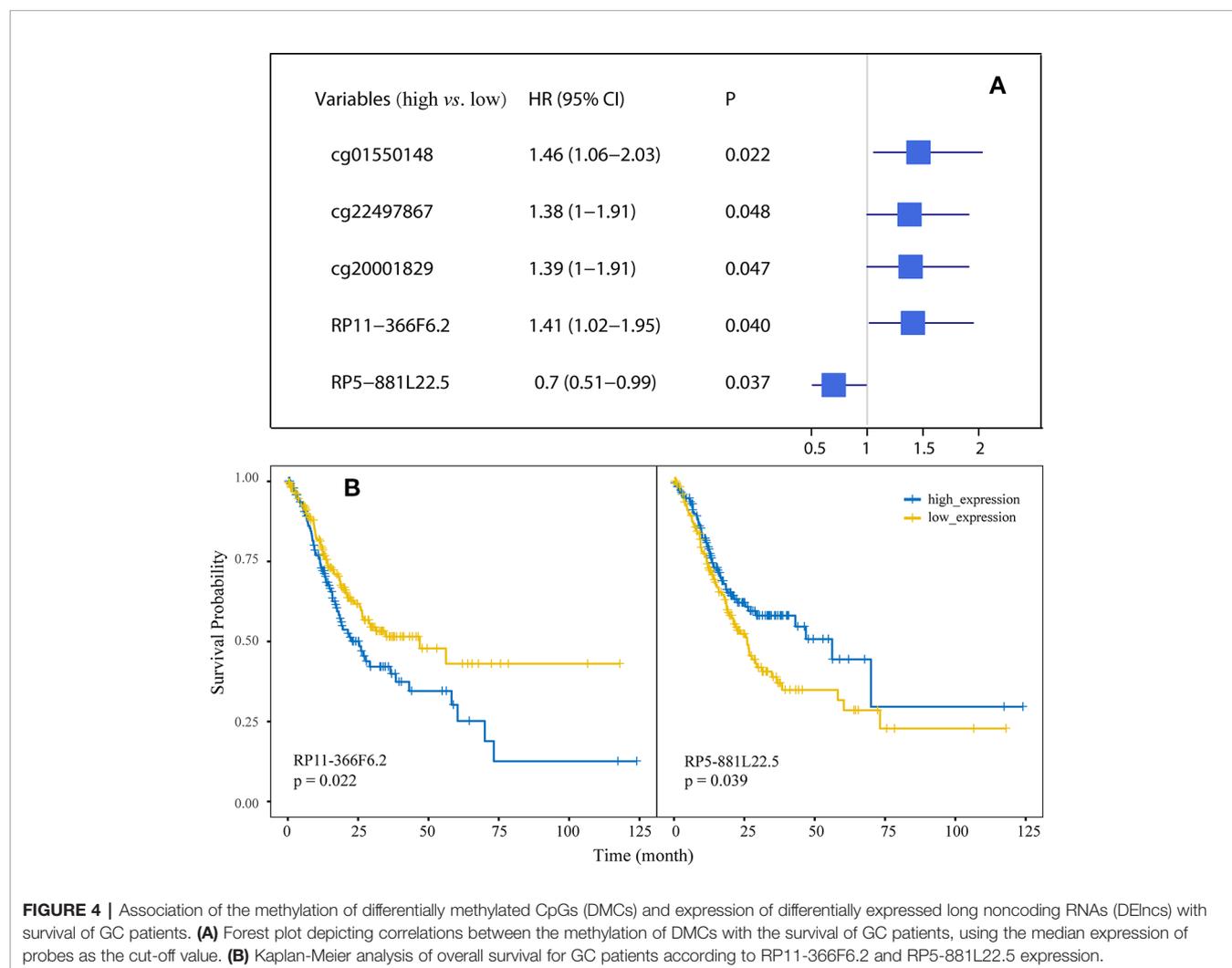

FIGURE 4 | Association of the methylation of differentially methylated CpGs (DMCs) and expression of differentially expressed long noncoding RNAs (DElncs) with survival of GC patients. **(A)** Forest plot depicting correlations between the methylation of DMCs with the survival of GC patients, using the median expression of probes as the cut-off value. **(B)** Kaplan-Meier analysis of overall survival for GC patients according to RP11-366F6.2 and RP5-881L22.5 expression.

TABLE 1 | Univariate and multivariate Cox regression analysis of variables associated with gastric cancer (GC) patient survival.

Variables	Univariate analysis			Multivariate analysis		
N=351	HR	95% CI	P	HR	95% CI	P
Age (≥67/ < 67)	1.44	1.04–2.00	0.029	1.49	1.06–2.11	0.023
Sex (male/female)	1.33	0.93–1.89	0.115	1.31	0.91–1.89	0.151
Tumor_stage (III+IV/I+II)	1.85	1.29–2.63	<0.001	1.89	1.32–2.70	<0.001
RP11-366F6.2 (high/low) *	1.41	1.02–1.95	0.040	1.45	1.03–2.05	0.033
RP5-881L22.5 (high/low) [†]	0.70	0.51–0.99	0.037	0.68	0.49–0.96	0.030

HR, hazard ratio; CI, confidence interval.

*Using the mean expression of genes as the cut-off value.

[†] Using the median expression of genes as the cut-off value.

ZNF667-AS1 were assigned GO molecular function, involving oxidoreductase activity, nucleotide diphosphatase activity, and transcriptional activator activity (**Figure 5A**). In addition, GSEA analysis was performed to identify the associated biological processes and signaling pathways for these deregulated lncRNAs. As an example, we explored a functionally unknown lncRNA, RP5-881L22.5, with significantly hypomethylated promoter regions that was upregulated in GC tissue and implicated with prognosis (**Figure 5B**). The expression of RP5-881L22.5 was positively correlated with "VECCHI_GASTRIC_CANCER_EARLY_UP" set, in which upregulated genes could differ early GC and normal tissue samples. The "NABA_ECM_GLYCOPROTEINS" set was enriched in the RP5-881L22.5 low expression group, which implied that this lncRNA could suppress tumor metastasis (**Figure 5C**).

DISCUSSION

Cancer involves a complex regulatory network, and therefore integrating multiple omics data is required in the era of precision medicine (Olivier et al., 2019). The increasing applications of multi-omic profiling of GC have delivered new insight into the dynamics of this cancer type. In this study, we characterized DNA methylation in the promoters of lncRNA-encoding genes and inferred the potential lncRNAs regulated by aberrant DNA methylation in GC. Differential analyses were performed to compare DNA methylation and gene expression patterns between GC and normal tissues, and 548 DMCs and 2,399 DElncs were obtained. We thus identified lncRNAs (such as HOTAIR, HOTTIP, HOXA11-AS, HOXB-AS4, and HOXC-AS3) that could be regulated by aberrant DNA methylation via combination analysis. We further divided the potentially epigenetic regulated lncRNAs into different groups to explore their biological and clinical relationships with GC and found that the expressions of RP11-366F6.2 and RP5-881L22.5 were related to the prognosis of GC.

Methylation that interferes with transcription machinery binding to DNA has been reported to be highly associated with repression of gene transcription (Suzuki and Bird, 2008). Our research showed that a large number of lncRNAs were epigenetically deregulated by promoter methylation, and lncRNAs were globally hypomethylated, which was consistent with the previous observation that increased global DNA hypomethylation was a major event for the development and progression of cancer (Zeng et al., 2017). HOX genes are a subset of evolutionarily conserved homeobox genes that encode a class of important transcription factors that function in numerous developmental processes (Quinonez and Innis, 2014). By characterizing the transcriptional landscape of the four human HOX loci (A–D) at five base pair resolution in 11 anatomic sites, Rinn et al. identified 231 HOX non-coding RNAs (Rinn et al., 2007). HOTAIR is located in the HOXC cluster and serves as a scaffold protein by binding Polycomb repressive complex 2 (PRC2, including SUZ12, EED, and EZH2) via its 5′-domain and the LSD1/CoREST/REST complex via its 3′-domain, mediating gene silencing and reprogramming the overall chromatin dynamics in GC (Qu et al., 2019). Like HOTAIR, HOXA11-AS recruits EZH2 along with the histone demethylase LSD1 or DNMT1, which promotes proliferation and invasion of GC (Sun et al., 2016). HOTTIP enhances the expression of neighboring HOXA genes, particularly HOXA13 (Chang et al., 2016). HOXC-AS3, an antisense transcript of HOXC10, mediates gene transcriptional regulation in the tumorigenesis of GC by binding to YBX1 (Zhang et al., 2018). Genome-wide screening isolated HOXB-AS4 as specifically methylated in pancreatic cancer cells, which was useful to assess a cancer cell fraction in DNA samples (Ishihara et al., 2018).

Therapeutic targets and prognosis prediction from a comprehensive analysis of multi-omics data and clinical profiles is a critical for better understanding the biological complexity of GC. We identified two hypomethylated DElncs (RP11-366F6.2 and RP5-881L22.5) in GC that were significantly associated with overall survival. RP11-366F6.2, also called MAGEA4-AS1, is located in chrX and was reported to have significantly high expression in several tumor tissues, such as breast cancer and laryngeal squamous cell carcinoma (Yuan et al., 2017; Liu and Ye, 2019). Although pre-ranked GSEA analysis for RP11-366F6.2 returned no significantly gene sets, examining the functional roles of deregulated genes (such as MAGEA4, MAGEA10, HOXD10 and IGF2BP1) in the leading edge set indicated RP11-366F6.2 might be associated with tumor invasion and metastasis (Suzuki et al., 2008; Schultz-Thater et al., 2011; Xu et al., 2019). Regarding RP5-881L22.5, Zhu et al. developed an eleven-lncRNA signature, including this lncRNA, which could provide an effective individual mortality risk prediction and risk stratification in GC patients (Zhu et al., 2018). However, the biological functions of RP5-881L22.5 have
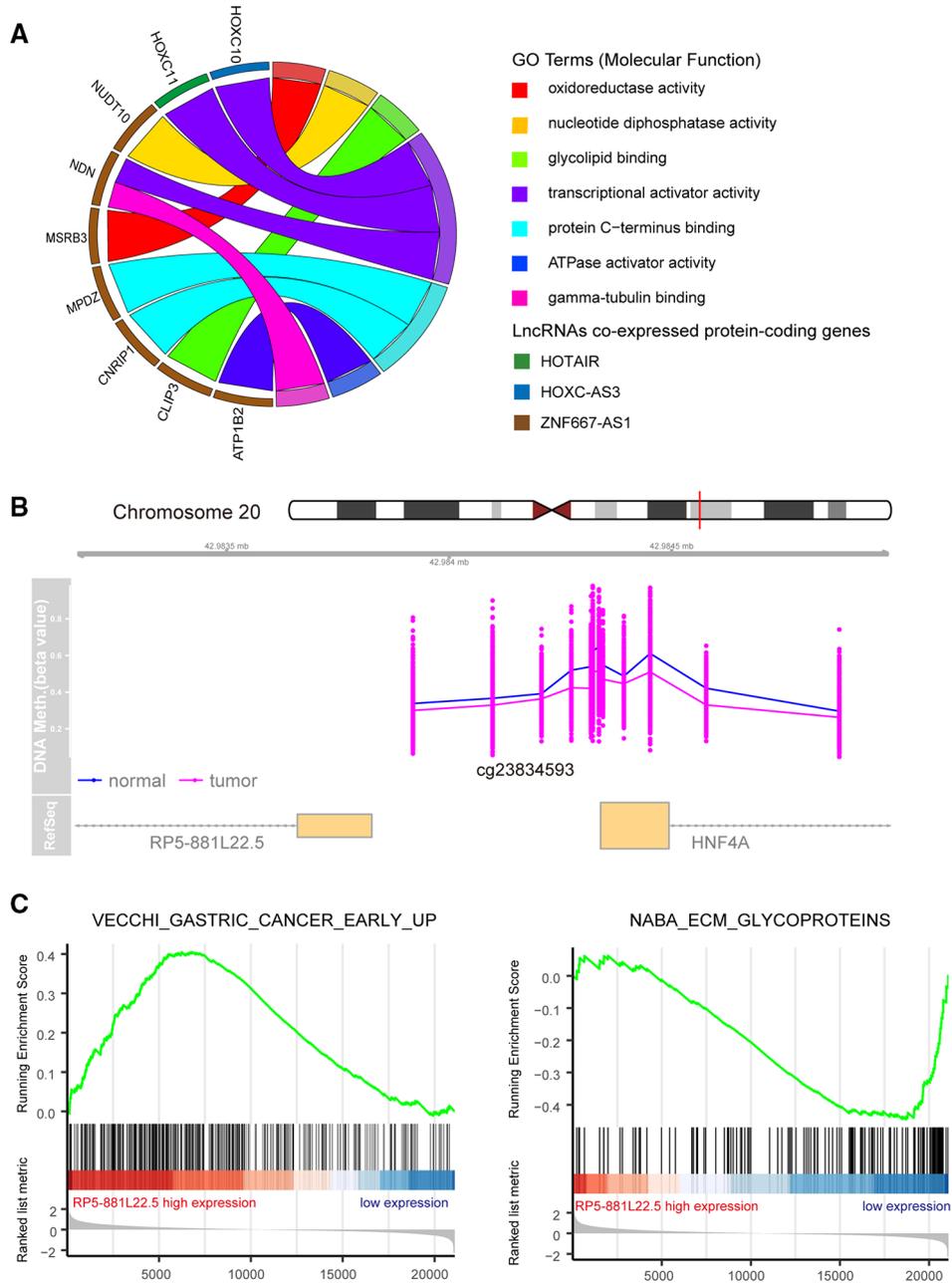
FIGURE 5 | Functional annotation for differentially methylated differentially expressed long noncoding RNAs (DElncs). **(A)** Circular plot of molecular function regulator Gene Ontology Term. **(B)** RP5-881L22.5 promoter hypermethylation in gastric cancer (GC) tumors compared with normal tissues. **(C)** Gene set enrichment analysis of RP5-881L22.5 in the The Cancer Genome Atlas (TCGA) dataset.

not been determined. GSEA results revealed that RP5-881L22.5 was likely to be involved in an extracellular matrix (ECM) interaction pathway. Glycoproteins make the ECM a cohesive network of molecules by linking cells together with structural components (Nallanthighal et al., 2019). Adhesive glycoproteins can bind to ECM components to activate downstream signaling pathways to regulate epithelial-mesenchymal transition, self-renewal, migration, differentiation, and proliferation (Naba

et al., 2016; Song et al., 2017). For example, the adhesion of cancer cells to fibronectin, a major adhesive ECM glycoprotein, remodels the tumor vasculature, enhances tumorigenicity, and facilitates metastasis. This mechanism could partly explain why decreased RP5-881L22.5 expression indicated a poor prognosis for GC patients.

To investigate the effect of epigenetically deregulated lncRNA in biological processes and pathways, an integrated analysis of

DElncs and predicted mRNAs expression was performed. GO analysis revealed that these lncRNAs were involved in dysregulated transcriptional programs that invariably lead to cancer. We also found that predicted mRNAs HOXC10 and MSRB3 in GO analysis were significantly associated with overall survival in GC patients (Kim et al., 2019; Ma et al., 2019).

Several limitations in the present study should be pointed out. First, integrated analysis of genome-wide DNA methylation and lncRNA expression was based on the 27K Illumina array platform, which only contained 27,578 individual registered probes, and thus some possibly important methylation differences may be lacking from the current results. Second, the results of the present study are preliminary and primarily derived from bioinformatics analysis, and lack functional validation of the epigenetically deregulated lncRNAs. Third, due to limited availability of clinical data, it was not possible to obtain deeper insights into characterizing phenotype-genotype relationships.

In conclusion, the present results provide evidence for the changes of widespread DNA methylation of lncRNA-encoding genes in GC patients. The candidate factors identified in this study might function as potential molecular phenotypic biomarkers, especially RP11-366F6.2 and RP5-881L22.5, which were associated with prognosis. Our results help elucidate a more detailed explanation of epigenetic mechanisms for GC and deepen our understanding of the aberrantly methylated patterns in lncRNA-encoding genes.

DATA AVAILABILITY STATEMENT

The following information was supplied regarding data availability:.

The level 3 TCGA data for DNA methylation arrays and lncRNA expression are available in Xena website.

HumanMethylation27: https://gdc.xenahubs.net/download/TCGA-STAD.methylation27.tsv.gz

HumanMethylation450: https://gdc.xenahubs.net/download/TCGA-STAD.methylation450.tsv.gz

HTSeq-Counts: https://gdc.xenahubs.net/download/TCGA-STAD.htseq_counts.tsv.gz

HTSeq-FPKM-UQ: https://gdc.xenahubs.net/download/TCGA-STAD.htseq_fpkm-uq.tsv.gz

Phenotype: https://gdc.xenahubs.net/download/TCGA-STAD.GDC_phenotype.tsv.gz

Survival data: https://gdc.xenahubs.net/download/TCGA-STAD.survival.tsv.gz

The microarray data that support this study are available through the NCBI database under accession GSE30601.

The data used to support the findings of this study are included within the article.

AUTHOR CONTRIBUTIONS

PS and WG designed the study. PS and LW collected, analysed and interpreted the data. PS and LW wrote the draft. PS and WG edited the manuscript.

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00091/full#supplementary-material

REFERENCES

Benjamini, Y., Drai, D., Elmer, G., Kafkafi, N., and Golani, I. (2001). Controlling the false discovery rate in behavior genetics research. Behav. Brain Res. 125 (1-2), 279–284. doi: 10.1016/s0166-4328(01)00297-2

Cancer Genome Atlas Research Network (2014). Comprehensive molecular characterization of gastric adenocarcinoma. Nature 513 (7517), 202–209. doi: 10.1038/nature13480

Chang, S., Liu, J., Guo, S., He, S., Qiu, G., Lu, J., et al. (2016). HOTTIP and HOXA13 are oncogenes associated with gastric cancer progression. Oncol. Rep. 35 (6), 3577–3585. doi: 10.3892/or.2016.4743

Fortin, J.-P., Triche, T. J., and Hansen, K. D. (2017). Preprocessing, normalization and integration of the Illumina HumanMethylationEPIC array with minfi. Bioinf. (Oxford England) 33 (4), 558–560. doi: 10.1093/bioinformatics/btw691

Guo, W., Lv, P., Liu, S., Xu, F., Guo, Y., Shen, S., et al. (2018). Aberrant methylation-mediated downregulation of long noncoding RNA C5orf66-AS1 promotes the development of gastric cardia adenocarcinoma. Mol. Carcinog. 57 (7), 854–865. doi: 10.1002/mc.22806

Honma, R., Goto, K., Sakamoto, N., Sekino, Y., Sentani, K., Oue, N., et al. (2017). Expression and function of Uc.160+, a transcribed ultraconserved region, in gastric cancer. Gastric Cancer 20 (6), 960–969. doi: 10.1007/s10120-017-0714-9

Ishihara, H., Yamashita, S., Amano, R., Kimura, K., Hirakawa, K., Ueda, T., et al. (2018). Pancreatic cancer cell fraction estimation in a DNA sample. Oncology 95 (6), 370–379. doi: 10.1159/000491637

Kim, J., Bae, D. H., Kim, J. H., Song, K. S., Kim, Y. S., and Kim, S. Y. (2019). HOXC10 overexpression promotes cell proliferation and migration in gastric cancer. Oncol. Rep. 42 (1), 202–212. doi: 10.3892/or.2019.7164

Kottorou, A., Antonacopoulou, A., Dimitrakopoulos, F.-I., Diamantopoulou, G., Theodorakopoulos, T., Oikonomou, C., et al. (2016). Diagnostic value of methylation status of T-UCRs for colorectal cancer. Ann. Oncol. 27 (suppl_6), vi188. doi: 10.1093/annonc/mdw370.118

Liu, Y., and Ye, F. (2019). Construction and integrated analysis of crosstalking ceRNAs networks in laryngeal squamous cell carcinoma. PeerJ 7, e7380. doi: 10.7717/peerj.7380

Liu, Y. W., Sun, M., Xia, R., Zhang, E. B., Liu, X. H., Zhang, Z. H., et al. (2015). LincHOTAIR epigenetically silences miR34a by binding to PRC2 to promote the epithelial-to-mesenchymal transition in human gastric cancer. Cell Death Dis. 6, e1802. doi: 10.1038/cddis.2015.150

Liu, H.-T., Liu, S., Liu, L., Ma, R.-R., and Gao, P. (2018). EGR1-mediated transcription of lncRNA-HNF1A-AS1 promotes cell-cycle progression in gastric cancer. Cancer Res. 78 (20), 5877–5890. doi: 10.1158/0008-5472.CAN-18-1011

Ma, X., Wang, J., Zhao, M., Huang, H., and Wu, J. (2019). Increased expression of methionine sulfoxide reductases B3 is associated with poor prognosis in gastric cancer. *Oncol. Lett.* 18 (1), 465–471. doi: 10.3892/ol.2019.10318

Mattick, J. S., and Rinn, J. L. (2015). Discovery and annotation of long noncoding RNAs. *Nat. Struct. Mol. Biol.* 22 (1), 5–7. doi: 10.1038/nsmb.2942

McCarthy, D. J., Chen, Y., and Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res.* 40 (10), 4288–4297. doi: 10.1093/nar/gks042

Mosquera Orgueira, A. (2015). Hidden among the crowd: differential DNA methylation-expression correlations in cancer occur at important oncogenic pathways. *Front. In Genet.* 6, 163. doi: 10.3389/fgene.2015.00163

Naba, A., Clauser, K. R., Ding, H., Whittaker, C. A., Carr, S. A., and Hynes, R. O. (2016). The extracellular matrix: tools and insights for the "omics" era. *Matrix Biol.: J. Int. Soc. Matrix Biol.* 49, 10–24. doi: 10.1016/j.matbio.2015.06.003

Nallanthighal, S., Heiserman, J. P., and Cheon, D.-J. (2019). The role of the extracellular matrix in cancer stemness. *Front. In Cell Dev. Biol.* 7, 86. doi: 10.3389/fcell.2019.00086

Olivier, M., Asmis, R., Hawkins, G. A., Howard, T. D., and Cox, L. A. (2019). The need for multi-omics biomarker signatures in precision medicine. *Int. J. Mol. Sci.* 20 (19), 4781. doi: 10.3390/ijms20194781

Qu, X., Alsager, S., Zhuo, Y., and Shan, B. (2019). HOX transcript antisense RNA (HOTAIR) in cancer. *Cancer Lett.* 454, 90–97. doi: 10.1016/j.canlet.2019.04.016

Quinonez, S. C., and Innis, J. W. (2014). Human HOX gene disorders. *Mol. Genet. Metab.* 111 (1), 4–15. doi: 10.1016/j.ymgme.2013.10.012

Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., et al. (2007). Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell* 129 (7), 1311–1323. doi: 10.1016/j.cell.2007.05.022

Schultz-Thater, E., Piscuoglio, S., Iezzi, G., Le Magnen, C., Zajac, P., Carafa, V., et al. (2011). MAGE-A10 is a nuclear protein frequently expressed in high percentages of tumor cells in lung, skin and urothelial malignancies. *Int. J. Cancer* 129 (5), 1137–1148. doi: 10.1002/ijc.25777

Shahabi, S., Kumaran, V., Castillo, J., Cong, Z., Nandagopal, G., Mullen, D. J., et al. (2019). LINC00261 is an epigenetically regulated tumor suppressor essential for activation of the DNA damage response. *Cancer Res.* 79 (12), 3050–3062. doi: 10.1158/0008-5472.can-18-2034

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13 (11), 2498–2504. doi: 10.1101/gr.1239303

Song, P., Jiang, B., Liu, Z., Ding, J., Liu, S., and Guan, W. (2017). A three-lncRNA expression signature associated with the prognosis of gastric cancer patients. *Cancer Med.* 6 (6), 1154–1164. doi: 10.1002/cam4.1047

Song, Y., Wang, R., Li, L.-W., Liu, X., Wang, Y.-F., Wang, Q.-X., et al. (2019). Long non-coding RNA HOTAIR mediates the switching of histone H3 lysine 27 acetylation to methylation to promote epithelial-to-mesenchymal transition in gastric cancer. *Int. J. Oncol.* 54 (1), 77–86. doi: 10.3892/ijo.2018.4625

Spizzo, R., Almeida, M. I., Colombatti, A., and Calin, G. A. (2012). Long non-coding RNAs and cancer: a new frontier of translational research? *Oncogene* 31 (43), 4577–4587. doi: 10.1038/onc.2011.621

Sun, M., Nie, F., Wang, Y., Zhang, Z., Hou, J., He, D., et al. (2016). LncRNA HOXA11-AS promotes proliferation and invasion of gastric cancer by scaffolding the chromatin modification factors PRC2, LSD1, and DNMT1. *Cancer Res.* 76 (21), 6299–6310. doi: 10.1158/0008-5472.can-16-0356

Suzuki, M. M., and Bird, A. (2008). DNA methylation landscapes: provocative insights from epigenomics. *Nat. Rev. Genet.* 9 (6), 465–476. doi: 10.1038/nrg2341

Suzuki, S., Sasajima, K., Sato, Y., Watanabe, H., Matsutani, T., Iida, S., et al. (2008). MAGE-A protein and MAGE-A10 gene expressions in liver metastasis in patients with stomach cancer. *Br. J. Cancer* 99 (2), 350–356. doi: 10.1038/sj.bjc.6604476

Torre, L. A., Bray, F., Siegel, R. L., Ferlay, J., Lortet-Tieulent, J., and Jemal, A. (2015). Global cancer statistics 2012. *CA Cancer J. Clin.* 65 (2), 87–108. doi: 10.3322/caac.21262

Turner, S. D. (2018). qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. *J. Open Source Softw.* 3 (25), 731. doi: 10.21105/joss.00731

Wu, D.-C., Wang, S. S. W., Liu, C.-J., Wuputra, K., Kato, K., Lee, Y.-L., et al. (2017). Reprogramming antagonizes the oncogenicity of HOXA13-long noncoding RNA HOTTIP axis in gastric cancer cells. *Stem Cells (Dayton Ohio)* 35 (10), 2115–2128. doi: 10.1002/stem.2674

Xu, F., Li, C. H., Wong, C. H., Chen, G. G., Lai, P. B. S., Shao, S., et al. (2019). Genome-wide screening and functional analysis identifies tumor suppressor long noncoding RNAs epigenetically silenced in hepatocellular carcinoma. *Cancer Res.* 79 (7), 1305–1317. doi: 10.1158/0008-5472.CAN-18-1659

Yu, G., Wang, L. G., Han, Y., and He, Q. Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 16 (5), 284–287. doi: 10.1089/omi.2011.0118

Yuan, N., Zhang, G., Bie, F., Ma, M., Ma, Y., Jiang, X., et al. (2017). Integrative analysis of lncRNAs and miRNAs with coding RNAs associated with ceRNA crosstalk network in triple negative breast cancer. *Onco Targets Ther.* 10, 5883–5897. doi: 10.2147/ott.s149308

Zeng, X.-Q., Wang, J., and Chen, S.-Y. (2017). Methylation modification in gastric cancer and approaches to targeted epigenetic therapy (Review). *Int. J. Oncol.* 50 (6), 1921–1933. doi: 10.3892/ijo.2017.3981

Zhang, E.-b., Kong, R., Yin, ,.D.-d., You, ,.L.-h., Sun, M., Han, L., et al. (2014). Long noncoding RNA ANRIL indicates a poor prognosis of gastric cancer and promotes tumor growth by epigenetically silencing of miR-99a/miR-449a. *Oncotarget* 5 (8), 2276–2292. doi: 10.18632/oncotarget.1902

Zhang, E., He, X., Zhang, C., Su, J., Lu, X., Si, X., et al. (2018). A novel long noncoding RNA HOXC-AS3 mediates tumorigenesis of gastric cancer by binding to YBX1. *Genome Biol.* 19 (1), 154. doi: 10.1186/s13059-018-1523-0

Zhu, M., Wang, Q., Luo, Z., Liu, K., and Zhang, Z. (2018). Development and validation of a prognostic signature for preoperative prediction of overall survival in gastric cancer patients. *Onco Targets Ther.* 11, 8711–8722. doi: 10.2147/ott.s181741

Zhuo, W., Liu, Y., Li, S., Guo, D., Sun, Q., Jin, J., et al. (2019). Long noncoding RNA GMAN, up-regulated in gastric cancer tissues, is associated with metastasis in patients and promotes translation of Ephrin A1 by competitively binding GMAN-AS. *Gastroenterology* 156 (3), 676–691.e611. doi: 10.1053/j.gastro.2018.10.054

Zouridis, H., Deng, N., Ivanova, T., Zhu, Y., Wong, B., Huang, D., et al. (2012). Methylation subtypes and large-scale epigenetic alterations in gastric cancer. *Sci. Transl. Med.* 4 (156), 156ra140. doi: 10.1126/scitranslmed.3004504