Identification of dysregulated long noncoding RNA and associated mechanism in gastric cancer

Liang Shao
The people's hospital of Shouguang https://orcid.org/0000-0001-8612-647X
Yanan Ming
Ruprecht Karls Universitat Heidelberg Medizinische Fakultat Heidelberg
Zhaoming Zhou (zhouzhaoming08@outlook.com)

Research article

Keywords: Gastric Cancer, Lnc RNA, Biomarker, IGF2BP3, FOLR1

Posted Date: June 2nd, 2020

DOI: https://doi.org/10.21203/rs.3.rs-30651/v1

Abstract

Gastric cancer (GC) is one of the most malignant epithelial tumors. The incidence of GC varies worldwide, and nearly half of the cases occur in Asian countries, especially in Japan and China. GC is the second leading cause of cancer-related deaths in the world, and current prognosis of advanced GC remains dismal despite improvements in diagnosis and therapy. Our current study aimed to identify significant long noncoding RNAs (lncRNAs) that could be used for prognosis and for the elucidation of associated molecular mechanisms.

This study identified a substantial variety of reports using high-throughput lncRNA detection to investigate the expression of lncRNAs in GC development. However, the reported potentially diagnostic lncRNAs were unsuitable for meta-analysis, so we verified the expression of diagnostic lncRNAs and selected eight of them using Gene Expression Profiling Interactive Analysis (GEPIA). Next, we explored interactions of selected lncRNAs using Qiagen's IPA system. We also identified differentially expressed genes (DEGs) of GC using the GEO2R online tool and datasets GSE52149, GSE19826, and GSE79973. To reveal genes that could be regulated by the selected lncRNA in GC, we used a Venn diagram and selected IGF2BP3 and FOLR1 as potential downstream targets of lncRNAs H19 and PVT1, respectively. Expression of IGF2BP3 and FOLR1 in GC validated by GEPIA was related to the worse prognosis for GC patients as shown by Kaplan Meier plots. Considering that IGF2BP3 promotes the expression of H19 and PEG10, down-regulation of their expression may improve the prognosis for GC patients, although at this time there is no evidence for direct involvement of IGF2BP3 in the regulation of these lncRNAs. The second DEG, FOLR1, is a crucial component of cell metabolism and DNA synthesis/repair required for cancer cell division. However, the role of FOLR1 in the etiology and progression of GC requires further study.

In conclusion, using an integrated bioinformatic approach we identified eight significantly altered lncRNAs with diagnostic potential in GC patients. We also identified two axes - H19-IGF2BP3 and PVT1-FOLR1 – that may be related to the prognosis of GC and provide new insights into the etiology of GC and management of GC patients.

1 Introduction

Gastric cancer (GC) is one of the most common cancers worldwide, with an estimated 989,600 new cases and 738,000 deaths globally in 2008 [1, 2]. GC incidence varies significantly, from the highest rate in Eastern Europe, South America and East Asia to the lowest rate in North America. While risk factors for GC —e.g. diet, smoking, and chronic *H. pylori* infection—are unique for each region, approximately two-thirds of newly diagnosed patients have either locally advanced or metastatic disease. As a result, the 5-year overall survival of GC patients remains as low as 20%, and the median survival time after surgery is only 6 to 9 months for late-stage disease. Improved health care and screening programs in Japan have shown that over 70% of patients with early-stage of GC survive over 5 years [3], indicating a decisive role of early diagnosis and treatment for the survival of GC patients. The most common diagnostic methods - endoscopy and pathological examination - are the 'gold standard' for GC diagnosis. Unfortunately, they cannot be used for screening in many countries due to the costs of the procedure and potential risk of

patient's injury [4], while existing common serum markers—CEA, CA199, and CA72-4—lack sensitivity and specificity essential for early cancer screening.

Long noncoding RNA (lncRNA) is a 200-nucleotide long transcript that regulates gene expression and messenger RNA (mRNA) splicing in the nucleus. Over time, a growing number of studies explored the role of lncRNA in the regulation of different physiological and pathological functions, e.g. cell differentiation and proliferation [5], carcinogenesis [6], and metastasis [7]. Several lncRNAs have been identified as oncogenes and tumor-suppressors: for example, up-regulation of HOTAIR drives proliferation, migration, and invasion of GC cell [8], H19 has oncogenic activity in GC and colon cancer [9], while CASC2 suppresses proliferation of GC cells through the MAPK signaling pathway [10]. Recent evidence indicates that lncRNAs can also modulate and be regulated by cancer immune microenvironment [11], making lncRNAs a potential biomarker and a therapeutic target that can improve the management and treatment of GC.

In this study, we did a meta-analysis of lncRNAs to assess their overall accuracy for the diagnosis of GC. Using Gene Expression Profiling Interactive Analysis (GEPIA), we compared expression patterns in GC and normal tissue and found eight lncRNAs with marked differences in expression. We also identified two genes that had different levels of expression in normal tissue and GC and could interact with these lncRNAs. Taken together, our results suggest a connection between lncRNAs and prognosis in GC patients that in the future may serve as a potential diagnostic biomarker.

2 Materials And Methods

2.1 Search strategy and eligibility criteria

Publicly available databases (PubMed: https://www.ncbi.nlm.nih.gov/pubmed/; and EMBASE: https://www.embase.com) were comprehensively searched to identify relevant English-language articles reporting microarray data for human lncRNAs and published up to the end of 2018. The following keywords and phrases were used: (lncRNA OR long noncoding RNA) AND ((gastric cancer) OR GC OR stomach neoplasms OR (stomach AND neoplasms) OR (gastric AND cancer)). Duplicate articles were manually removed using Reference Manager (Thomson Reuters EndNote X7, New York, NY, USA). To determine eligible studies the titles, abstracts, and full texts were evaluated independently by two investigators. Another investigator extracted data from identified papers, and the reference lists of eligible articles were reviewed to obtain associated studies. All disagreements were resolved by an independent investigator. The criteria for inclusion were: 1) studies with a confirmed diagnosis of gastric cancer; 2) studies with lncRNAs microarray analysis and reports on altered lncRNAs; 3) studies on diagnostic value of lncRNAs in tissue, serum, plasma, peripheral blood, or gastric juice (if published data were sufficient to allow meta-analysis); 4) original articles published in English with full text available. Articles that did not satisfy these criteria were excluded. General information from the eligible studies was arranged in tables, and data on lncRNAs were pooled into forest plots.

2.2 mRNA microarray data information and processing DEGs

We used NCBI-GEO, an online public microarray database, to acquire gene expression profiles for GC and normal stomach tissues from GSE54129, GSE19826 and GSE79973 datasets that had been produced using GPL570 Platform ([HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array). GPL570 contained 111 human GC tissues and 21 noncancerous tissues, 12 adjacent normal/tumor-paired gastric tissues, and 10 pairs of GC tissue and adjacent non-tumor mucosa, respectively. DEGs in tumor and normal tissues were identified with GEO2R online tools [6]. The DEGs with |log2FC|< 0 were considered to be down-regulated, while the DEGs with |log2FC|> 0 - up-regulated. Three microarrays were compared, and upregulated and downregulated genes with |log2FC|> 2 and adjusted p-value less than 0.05 in TXT format were pooled into Venn software online to identify common DEGs.

2.3 Bioinformatics analysis

Bioinformatics analysis was performed to determine relationships between altered lncRNAs and DEGs of overlapping mRNA microarrays. Briefly, lncRNAs from selected studies were uploaded into GEPIA to validate their expression in GC tissues. Then lncRNAs with altered expression were imported into Qiagen's IPA system and overlaid with a global network of interactions in gastrointestinal disease. Next, we overlaid identified DEGs and genes related to the altered lncRNAs to find potential interactions. Finally, we used GEPIA and Kaplan Meier plotter online database to validate expression levels and effects of identified genes and lncRNAs on survival time.

2.4 Statistical methods

The confidence interval (CI) of diagnostic value was calculated using meta-disc (version 1.4; Ramony Cajal Hospital, Madrid, Spain), the results were considered significant for a two-sided p-value less than 0.05. Heterogeneity was inferred by calculating inconsistency (I2) (heterogeneity was considered substantial for I2 values above 50%), and the results were incorporated into a random-effects model. Potential reasons for heterogeneity were analyzed by regression analysis. The sensitivity and specificity of potential biomarkers were evaluated using summary receiver operating curve (sROC) and area under the curve (AUC). In addition, LR+ (positive likelihood ratio), LR- (negative likelihood ratio), and DOR (diagnostic odds ratio) were calculated.

3 Results

3.1 lncRNAs expression levels altered in GC

The summary of 13 studies reporting changes in lncRNA expression in GC tumorigenesis and development is presented in Table 1. Different types of samples were used: Zhang et al. compared lncRNA between GC and non-GC patients using tissue and plasma [12]; gastric mucosae were used in one study [13]; all other studies analyzed GC cancer tissue and adjacent healthy tissue. Among all identified studies, differences in lncRNAs expression levels were detected for at least 75 lncRNAs [12]. Because of the re-annotation of the published microarray database, two studies did not provide fold changes of dysregulated lncRNAs [12, 14, 15]. A study from Hu et al. used a 1.5-fold change for selection criteria [16], while 2 fold change was used for other studies [12, 17–24]. The median age of GC patients enrolled in the analysis was at least 57.8

years, except one study [23] that did not have detailed information. Gender distribution and histopathological information of GC patients are shown in Table 1.

Table 1
LncRNA expression profiles included in the systematic review

Studies	Platform of LcRNA microarray	Sources of samples (N or pairs)	The number of dysregulated LncRNA	Gastric cancer patients		
				Age	Sex (M/F)	AJCC stage
Cao, 2013[14]	the Affymetrix GeneChip Human Exon 1.0 ST Array (re-annotation)	22pairs, tissue	88(p < 0,01)	59.4(34–84)	53/27	4/7/54/15(I/II/III/IV)
Li, 2016[17]	RiboArray Custom Array1*90K + q RT-PCR	10/10, tissue	1046(folds > 2 & p < 0,01)	58.5(45–70)	7/3	1/6/3 (II/III/IV)
Zhang, 2017[12]	The LncRNA Human Gene Expression Microarray V4.0	15/15 plasma; tissue	Plasma:77 Tissue:75(folds > 2 & p < 0,05)	60.21	9/6	7/8 (II/III)
Hu, 2014[16]	ArrayStar lncRNA microarray	10/10, tissues	1368(folds > 1,5 & p < 0,05)	63(48–76)	8/2	NR
Song, 2016[18]	the Human LncRNA Expression Microarray V3.0	6/2, tissue	1379(folds > 2 & p < 0,05)	60.2(45–77)	4/2	2/4 (II/III)
Lin, 2014[19]	The Human 12 × 135 k lncRNA expression microarray	8 pairs, tissue	2621(folds > 2 & p < 0,05)	58.75	6/2	1/3/4(I/II/III)
Song, 2013[20]	NimbleGen Hybridization System	3/3, tissue	135(folds > 2)	73(55–88)	2/1	NR
Wang, 2014[21]	the Human LncRNA Expression Microarray V2.0	15/15, tissue	5139(folds > 2)	57.8(42–77)	13/2	1/1/13(I/II/III)

AJCC: American Joint Committee on Cancer; NR: No reported, qRT-PCR: Real-time Quantitative polymerase chain reaction; M: male; F: female.

Studies	Platform of LcRNA microarray	Sources of samples (N or pairs)	The number of dysregulated LncRNA	Gastric cancer patients		
				Age	Sex (M/F)	AJCC stage
Yuan, 2016[22]	Affymetrix Human Genome U133 Plus 2.0 chips	20 pairs ,tissue	2323(folds > 2 & p < 0,05)	62(41–76)	16/4	4/5/8/3(I/II/III/IV)
Tian, 2017[15]	Affymetrix Human Genome U133 Plus 2.0 chips	10 pairs ,tissue	339(p < 0,01)	57.9	7/3	2/5/3 (II/III/IV)
Gu, 2015[23]	Arraystar Human LncRNA Microarray v2.0	6/6, tissue	1297(folds > 2 & p < 0,05)	NR	NR	NR
Zhao, 2015[24]	Arraystar Human LncRNA Microarray v3.0	6/6, tissue	3141(folds > 2 & p < 0,05)	61.2(42–73)	3/3	3/3(II/III)
Liu, 2018 [47]	Agilent human lncRNA V6 Microarray	4/4, plasma	267 (p < 0,05)	61(29–83)	2/2	NR

AJCC: American Joint Committee on Cancer; NR: No reported, qRT-PCR: Real-time Quantitative polymerase chain reaction; M: male; F: female.

In summary, the analysis of identified microarray data showed substantial lncRNAs alteration in GC patients. However, data reported by different research groups were extremely variable. Thus, we concentrated on the potential of lncRNAs for GC diagnosis.

3.2 Meta-analysis of differentially expressed lncRNAs in GC patients.

To further investigate the diagnostic value of lncRNAs in GC, all articles exploring lncRNAs as a novel biomarker for GC patients were collected using the search strategy indicated in the flowchart in Fig. 1. Twenty-three studies were included and pooled into meta-analysis.

The number of patients in each study ranged from 30 to 132, and tissue, plasma, serum, or gastric juice samples were used. Besides β-actin or GAPDH, U6 [25, 26] and 18 s RNAs [27] were used as endogenous standards for diagnostic evaluation. Quantitative methods and cut-off values were also different in

different studies, and both individual lncRNAs and panels [27, 28] were selected as novel diagnostic biomarkers for GC. Additionally, classic GC biomarkers (e.g. CEA and CA19-9) were compared with novel GC biomarkers [28], and the lncRNA panel showed markedly higher AUC value for discriminating GC patients from controls. Considering this evidence, we performed a meta-analysis with meta-disc software version 1.4. We pooled data from various specimens and generated forest plots shown in Fig. 2. The pooled sensitivity was 0.76 (95% CI: 0.74–0.77; Q = 195.59, p = 0.0000, I2% = 85.2%) and the specificity was 0.66 (95% CI: 0.64–0.68; Q = 208.98, p = 0.0000, I2% =86.1%), which indicated a presence of substantial heterogeneity. Then a random-effects model was used to re-analyze the diagnostic threshold of pooled data. The Spearman correlation coefficient was 0.238 (p = 0.214, data not shown), suggesting no evidence of a diagnostic threshold. Afterward, forest plots of DOR were generated, which revealed that substantial heterogeneity was still present. This might result from the discrepancy of the studied populations, endogenous references, or specimen types. Meta-regression analysis on the possible factors indicated that specimen type was probably the reason for heterogeneity. Thus, the results (e.g. sensitivity) extracted from identified studies could not be simply pooled and were only suitable for subgroup analyses. Filtering studies based on specimen type reduced heterogeneity, however, it was still higher than acceptable levels. On the sROC curve of plasma samples, which included 16 LncRNAs, the maximum joint sensitivity and specificity (Q value) was 0.7443, and the area under the curve was 0.8096, indicating a moderate level of overall accuracy. The combined sensitivity, specificity, LR+, LR- and DOR in plasma were 0.84 (95% CI: 0.81–0.86; I2% =82.1%), 0.50 (95% CI: 0.56–0.62; I2% =86.8%), 2.32 (95% CI: 1.88–2.85; I2% =84.4%), 0.27 (95% CI: 0.20–0.36; I2% =73.7%), and 9.53 (95% CI: 6.21–14.61; I2% =68.4%), respectively (**Fig. S1**). Thus, the results indicated that a pooled study was not appropriate. In the other 14 studies that used tissues as specimens the pooled sensitivity, specificity, LR+, LR-, and DOR were 0.69 (95% CI: 0.66–0.71; I2% =68.1%), 0.72 (95% CI: 0.69–0.74; I2% =76.9%), 2.50 (95% CI: 2.11–2.96; I2% =63.5%), 0.44 (95% CI: 0.38–0.50; I2% =52.2%), and 6.08 (95% CI: 4.65–7.96; I2% =51.8%), respectively (**Fig. S2**). The data showed a lower pooled sensitivity for tissue.

3.3 Validation of lncRNAs expression by GEPIA

Since heterogeneity was not reduced to an acceptable level through subgroup analysis, we further validated the expression of those lncRNAs between tumor tissues and normal tissues through GEPIA analysis. The analysis revealed that six lncRNAs (ABHD11-AS1, H19, PVT1, UCA, HOTTIP, and SUMO1) were reported to have significantly higher expression, while two lncRNAs (FER1L4 and LINC00982) showed noticeably lower expression in GCs samples compared to normal GC tissue. The expression of those lncRNAs is shown in Fig. 3.

3.4 Identification of DEGs in GCs and investigation of correlation to modulated lncRNAs by IPA

To investigate the underlying mechanism related to lncRNAs, we extracted 3944, 629 and 1406 DEGs from GSE52149, GSE19826 and GSE 79973 via GEO2R online tools, respectively. Venn diagram software was used to identify common DEGs in these datasets. A total of 226 common DEGs were identified, including

142 up-regulated genes ($p < 0.05$ & $log2FC < 2$) and 84 down-regulated genes ($p < 0.05$ & $log2FC > 2$) in the GC tissues (Fig. 4). Meanwhile, IPA analysis identified molecules that interacted with the altered lncRNAs in gastrointestinal diseases (Table 2). Then we pooled these molecules and DEGs in GC into the Venn diagram and identified two genes (IGF2BP3 and FOLR1), which probably interacted with altered lncRNAs in GC.

Table 2
The molecules interacted with selected eight LncRNAs through IPA system.

LncRNA	Related molecules
H19	CDKN3 Symbol ERK1/2 ATP7B staurosporine ABCB4 Gsk3 Map3k7 DMD CTCF lipopolysaccharide IGF2 SUZ12 carbon tetrachloride HNRNPA2B1 TERF2IP SOX9 Rb halofuginone PHB IL1 dicarbethoxydihydrocollidine DDX43 COL2A1 PDGF BB NOS2 TNF INSR tamoxifen PI3K (complex) CDC73 IGF2BP3 SOX2 APP TERC corticosterone H19 bucladesine EOMES IGF1R DICER1 Gm21596/Hmgb1 IRS1 DNMT3A IL1B beta-estradiol PGR E2F1 tetradecanoylphorbol acetate PARP1 ZBTB7A cigarette smoke ESR1 ERBB2 HNRNPU ZFP57
UCA1	Akt NUPR1 PTEN HNF1A ARID1A YAP1 TGFB1 ARL2 CEBPA CCND1 EZH2
PVT1	FOLR1 HR let-7 STAT5B NFIB LIN28A GAS2L3
FER1L4	RB1

ABCB4: ATP-binding cassette (ABC) subfamily B member 4; APP: Amyloid beta precursor protein; ARID1A: AT-rich interactive domain-containing protein 1A; ARL2: ADP Ribosylation Factor Like GTPase 2; ATP7B: ATPase copper transporting beta; CCND1: Cyclin D1; CDC73: Cell division cycle 73; CDKN3: Cyclin dependent kinase inhibitor 3; CEBPA: CCAAT enhancer binding protein alpha; COL2A1: Collagen type II alpha 1 chain; CTCF: CCCTC-binding factor; DDX43: DEAD-Box helicase 43;DMD: Dystrophin; DNMT3A: DNA methyltransferase; E2F1: E2F Transcription Factor 1; EOMES: Eomesodermin; ERK: extracellular signal-regulated kinase; ERBB2: Receptor tyrosine-protein kinase erbB-2; ESR1: Estrogen Receptor 1; EZH2: Enhancer of zeste homolog 2; FOLR1: Folate receptor 1; GAS2L3: Growth arrest specific 2 like 3; HNF1A: HNF1 homeobox A; HNRNPA2B1: Heterogeneous nuclear ribonucleoprotein A2/B1; HNRNPU: Heterogeneous nuclear ribonucleoprotein U; HR: HR Lysine demethylase and nuclear receptor corepressor; IGF2: Insulin-like growth factor 2; IGF2BP3: Insulin-like growth factor 2 mRNA binding protein 3; IGF1R: Insulin-like Growth Factor 1 Receptor; IL1: Interleukin-1; INSR: Insulin receptor; IRS1: Insulin Receptor Substrate 1; LIN28A: Lin-28 homolog A; NFIB: Nuclear factor 1 B; NOS2: Nitric oxide synthase 2; NUPR1: Nuclear protein 1, transcriptional regulator; PARP1: Poly(ADP-ribose) polymerase 1; PDGF BB: Platelet-derived growth factor with two B subunits; PGR: Progesterone receptor; : Prohibitin; PI3K: Phosphatidylinositol-4,5-bisphosphate 3-kinase; PTEN: Phosphatase and tensin homolog; RB1: The retinoblastoma protein; SOX2: SRY-Box Transcription Factor 2; SOX9: SRY-Box transcription factor 9; STAT5B: Signal transducer and activator of transcription 5B; TERF2IP: TERF2 interacting protein; TGFB1: Transforming growth factor beta 1; TNF: Tumor necrosis factor; TERC: Telomerase RNA component; YAP1: yes-associated protein 1; ZBTB7A: Zinc Finger And BTB Domain Containing 7A; ZFP57: Zinc Finger Protein57;

3.5 Validation of genes interacted with lncRNAs in GC

To investigate the potential role of these genes in GC, we further validated the expression of IGF2BP3 and FOLR1. GEPIA website and Kaplan Meier plotter (http://kmplot.com/analysis) were used to recognise the correlation between the expression of those genes and the prognosis of GC patients. We found a dramatically increased expression of IGF2BP3 and significantly reduced expression of FOLR1 in GC

patients compared to healthy controls (shown in **Fig. 5**). Meanwhile, the change of expression level for these two genes was reported to be correlated with worse survival time of GC patients (shown in **Fig. 5**).

4 Discussion

To improve the survival time of GC patients, early diagnosis and treatment have been recognised as effective methods. Thus, the exploration of useful biomarkers for early diagnosis and positive management based on the mechanism of GC development is required. To date, several available biomarkers, such as CA-199, CA72-4, and CEA, are used. However, the sensitivity and specificity of those biomarkers are low. The first study of lncRNAs in GC has been reported in 1997 [29], with more research focused on the clinical value of lncRNAs in GC diagnosis. Exploration of dysregulated lncRNAs as biomarkers for GC diagnosis has several advantages: 1) lncRNAs can be detected and resist ribonuclease degradation in body fluids [30]; 2) expression of lncRNAs has temporal and tissue specificity [31]; 3) ectopic expression of lncRNAs is responsible for tumorigenesis [32, 33]. Therefore, investigation of lncRNAs might produce novel diagnostic and prognostic biomarkers for GC and help us understand the molecular mechanisms of GC development and progression.

To explore the potential role of lncRNAs in GC, the present study reviewed and analyzed published studies that reported differentially expressed lncRNAs between GC and normal tissue using microarray analysis. Since there was a substantial variety of reported data, for meta-analysis we retrieved articles reporting on the diagnostic value of lncRNAs. However, the results indicated that the data pooled from all studies showed marked heterogeneity that was most likely associated with specimen types as shown by meta-regression analysis. Subsequently, we performed a subgroup meta-analysis. Although it revealed that individual or specific lncRNA combinations could potentially serve as novel biomarkers for diagnosis of GC, the heterogeneity was still too high. Meanwhile, we found that the data from different research groups had significant differences in quality. Through GEPIA analysis, we validated eight lncRNAs with significant differences of expression in GC compared to normal tissue. To explore the potential function of these lncRNAs in GC, we utilised IPA analysis to investigate molecules interacting with these lncRNAs in gastrointestinal diseases. We used bioinformatics methods based on three datasets (GSE52149, GSE19826, and GSE79973) to identify DEGs in GC. Then results for lncRNAs and DEGs were pooled into the Venn diagram to identify two genes (IGF2BP3 and FOLR1), which might be regulated by altered lncRNAs in GC samples. Finally, we utilised GEPIA and Kaplan Meier plotter analysis to validate that IGF2BP3 and FOLR1 both changed significantly and correlated with worse survival time in GC patients. Nevertheless, further research is still required to investigate mechanisms underlying these observations that might provide new effective treatments and improve outcomes for GC patients.

Insulin-like growth factor-2 mRNA-binding protein 3 (IGF2BP3), also known as IMP3, belongs to a conserved IGF2 mRNA-binding protein family. It has been first recognised in 1997, due to its high expression in pancreatic carcinoma [34]. Subsequently, IGF2BP3 has been found to be overexpressed in various tumors [35–38]. Moreover, it has been demonstrated to modulate tumor cell fate by promoting tumor growth [39], cell proliferation [2], drug-resistance [40], and invasiveness [41]. The expression of

IGF2BP3 has been shown to correlate with prognosis and metastasis of human cancer. Another study has shown that H19, PEG10, and IGF2BP3 promote expression of each other and that down-regulation of their expression can decrease cell proliferation, anchorage-independent growth, invasion, and chemoresistance in GC [42].

Folate receptor 1 (FOLR1) is a membrane-bound protein with a high affinity to folate that binds and transports folate with physiological levels into cells. Folate, one of the crucial components of cell metabolism and DNA synthesis and repair, is a requirement for the rapid division of cancer cells [43]. A higher expression of FOLR1 has been found in specific epithelial-derived malignant tumors compared to normal tissues [44] and has been illustrated to positively correlate with tumor grade and stage [43]. During early carcinogenesis, FOLR1 may promote cells to increase folate uptake and repair DNA damage [45]. Recently, FOLR1 has been confirmed as a potential target for immunotherapy with chimeric antigen receptor (CAR) T cell in GC [46]. However, at this time the role of FOLR1 in etiology and progression of GC is not fully understood.

Numerous studies have proven lncRNAs play a critical role in DEGs function in GC. However, as a regulator for various signaling pathways, very few studies report on the mechanism underlying these changes in lncRNAs and DEGs in GC. Therefore, the results of the present study might provide useful information that can guide researchers in studies of mechanisms involved and explorations of novel potential candidates as biomarkers for diagnosis of GC.

Declarations

Funding

Funding was not received.

Author’s contributions:

Shao and Ming analyzed and interpreted the data from eligible publications. Shao analyzed the GEO datasets and identified the potential interaction between selected LncRNAs and DEGs in GC patients. Zhou and Ming were major contributor in conclusion and writing the manuscript. All authors read and approved the final manuscript.

Acknowledgements:

Not applicable. The authors declare that they have no competing interests, and all authors should confirm its accuracy.”

Availability of data and the materials

The datasets used and/or analyzed during the current study are available from the GEO datasets.

Acknowledgement

None.

Conflict of Interst

None.

References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. International journal of cancer. 2010;127(12):2893–917.
2. Rivera Vargas T, Boudoukha S, Simon A, Souidi M, Cuvellier S, Pinna G, Polesskaya A. Post-transcriptional regulation of cyclins D1, D3 and G1 and proliferation of human cancer cells depend on IMP-3 nuclear localization. Oncogene. 2014;33(22):2866–75.
3. Isobe Y, Nashimoto A, Akazawa K, Oda I, Hayashi K, Miyashiro I, Katai H, Tsujitani S, Kodera Y, Seto Y, et al: **Gastric cancer treatment in Japan**: **2008 annual report of the JGCA nationwide registry**. *Gastric Cancer* 2011, **14**(4):301–316.
4. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. Group Q-: **QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies**. Ann Intern Med. 2011;155(8):529–36.
5. Chen J, Liu S, Hu X. Long non-coding RNAs: crucial regulators of gastrointestinal cancer cell proliferation. Cell death discovery. 2018;4:50.
6. Zhang J, Zhang P, Wang L, Piao HL, Ma L. Long non-coding RNA HOTAIR in carcinogenesis and metastasis. Acta Biochim Biophys Sin. 2014;46(1):1–5.
7. Serviss JT, Johnsson P, Grander D. An emerging role for long non-coding RNAs in cancer metastasis. Frontiers in genetics. 2014;5:234.
8. Liu XH, Sun M, Nie FQ, Ge YB, Zhang EB, Yin DD, Kong R, Xia R, Lu KH, Li JH, et al. Lnc RNA HOTAIR functions as a competing endogenous RNA to regulate HER2 expression by sponging miR-331-3p in gastric cancer. Mol Cancer. 2014;13:92.
9. Li H, Yu B, Li J, Su L, Yan M, Zhu Z, Liu B. Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. Oncotarget. 2014;5(8):2318–29.
10. Li P, Xue WJ, Feng Y, Mao QS. Long non-coding RNA CASC2 suppresses the proliferation of gastric cancer cells by regulating the MAPK signaling pathway. American journal of translational research. 2016;8(8):3522–9.

11. Luo X, Qiu Y, Jiang Y, Chen F, Jiang L, Zhou Y, Dan H, Zeng X, Lei YL, Chen Q. Long non-coding RNA implicated in the invasion and metastasis of head and neck cancer: possible function and mechanisms. Mol Cancer. 2018;17(1):14.
12. Zhang K, Shi H, Xi H, Wu X, Cui J, Gao Y, Liang W, Hu C, Liu Y, Li J, et al. Genome-Wide lncRNA Microarray Profiling Identifies Novel Circulating lncRNAs for Detection of Gastric Cancer. Theranostics. 2017;7(1):213–27.
13. Ren W, Zhang J, Li W, Li Z, Hu S, Suo J, Ying X. A Tumor-Specific Prognostic Long Non-Coding RNA Signature in Gastric Cancer. Medical science monitor: international medical journal of experimental clinical research. 2016;22:3647–57.
14. Cao WJ, Wu HL, He BS, Zhang YS, Zhang ZY. Analysis of long non-coding RNA expression profiles in gastric cancer. World journal of gastroenterology. 2013;19(23):3658–64.
15. Tian X, Zhu X, Yan T, Yu C, Shen C, Hong J, Chen H, Fang JY. Differentially Expressed lncRNAs in Gastric Cancer Patients: A Potential Biomarker for Gastric Cancer Prognosis. J Cancer. 2017;8(13):2575–86.
16. Hu Y, Wang J, Qian J, Kong X, Tang J, Wang Y, Chen H, Hong J, Zou W, Chen Y, et al. Long noncoding RNA GAPLINC regulates CD44-dependent cell invasiveness and associates with poor prognosis of gastric cancer. Cancer research. 2014;74(23):6890–902.
17. Li C, Liang G, Yao W, Sui J, Shen X, Zhang Y, Ma S, Ye Y, Zhang Z, Zhang W, et al. Differential expression profiles of long non-coding RNAs reveal potential biomarkers for identification of human gastric cancer. Oncol Rep. 2016;35(3):1529–40.
18. Song W, Liu YY, Peng JJ, Liang HH, Chen HY, Chen JH, He WL, Xu JB, Cai SR, He YL. Identification of differentially expressed signatures of long non-coding RNAs associated with different metastatic potentials in gastric cancer. Journal of gastroenterology. 2016;51(2):119–29.
19. Lin XC, Zhu Y, Chen WB, Lin LW, Chen DH, Huang JR, Pan K, Lin Y, Wu BT, Dai Y, et al. Integrated analysis of long non-coding RNAs and mRNA expression profiles reveals the potential role of lncRNAs in gastric cancer pathogenesis. Int J Oncol. 2014;45(2):619–28.
20. Song H, Sun W, Ye G, Ding X, Liu Z, Zhang S, Xia T, Xiao B, Xi Y, Guo J. Long non-coding RNA expression profile in human gastric cancer and its clinical significances. Journal of translational medicine. 2013;11:225.
21. Wang Y, Gao S, Liu G, Jia R, Fan D, Feng X. Microarray expression profile analysis of long non-coding RNAs in human gastric cardiac adenocarcinoma. Cellular physiology biochemistry: international journal of experimental cellular physiology biochemistry pharmacology. 2014;33(4):1225–38.
22. Yuan J, Yue H, Zhang M, Luo J, Liu L, Wu W, Xiao T, Chen X, Chen X, Zhang D, et al. Transcriptional profiling analysis and functional prediction of long noncoding RNAs in cancer. Oncotarget. 2016;7(7):8131–42.
23. Gu W, Gao T, Sun Y, Zheng X, Wang J, Ma J, Hu X, Li J, Hu M. LncRNA expression profile reveals the potential role of lncRNAs in gastric carcinogenesis. Cancer Biomark A. 2015;15(3):249–58.

24. Zhao J, Liu Y, Zhang W, Zhou Z, Wu J, Cui P, Zhang Y, Huang G. Long non-coding RNA Linc00152 is involved in cell cycle arrest, apoptosis, epithelial to mesenchymal transition, cell migration and invasion in gastric cancer. Cell cycle. 2015;14(19):3112–23.
25. Fu M, Huang Z, Zang X, Pan L, Liang W, Chen J, Qian H, Xu W, Jiang P, Zhang X. **Long noncoding RNA LINC00978 promotes cancer growth and acts as a diagnostic biomarker in gastric cancer**. *Cell Prolif* 2018, 51(1).
26. Pan L, Liang W, Gu J, Zang X, Huang Z, Shi H, Chen J, Fu M, Zhang P, Xiao X, et al. Long noncoding RNA DANCR is activated by SALL4 and promotes the proliferation and invasion of gastric cancer cells. Oncotarget. 2018;9(2):1915–30.
27. Ke D, Li H, Zhang Y, An Y, Fu H, Fang X, Zheng X. The combination of circulating long noncoding RNAs AK001058, INHBA-AS1, MIR4435-2HG, and CEBPA-AS1 fragments in plasma serve as diagnostic markers for gastric cancer. Oncotarget. 2017;8(13):21516–25.
28. Dong L, Qi P, Xu MD, Ni SJ, Huang D, Xu QH, Weng WW, Tan C, Sheng WQ, Zhou XY, et al. Circulating CUDR, LSINCT-5 and PTENP1 long noncoding RNAs in sera distinguish patients with gastric cancer from healthy controls. International journal of cancer. 2015;137(5):1128–35.
29. Wu MS, Wang HP, Lin CC, Sheu JC, Shun CT, Lee WJ, Lin JT. Loss of imprinting and overexpression of IGF2 gene in gastric adenocarcinoma. Cancer letters. 1997;120(1):9–14.
30. Shi T, Gao G, Cao Y. Long Noncoding RNAs as Novel Biomarkers Have a Promising Future in Cancer Diagnostics. Disease markers. 2016;2016:9085195.
31. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. Nature reviews Genetics. 2009;10(3):155–9.
32. Qiu MT, Hu JW, Yin R, Xu L. Long noncoding RNA: an emerging paradigm of cancer research. Tumour biology: the journal of the International Society for Oncodevelopmental Biology Medicine. 2013;34(2):613–20.
33. Zhang H, Chen Z, Wang X, Huang Z, He Z, Chen Y. Long non-coding RNA: a new player in cancer. J Hematol Oncol. 2013;6:37.
34. Mueller-Pillasch F, Lacher U, Wallrapp C, Micha A, Zimmerhackl F, Hameister H, Varga G, Friess H, Buchler M, Beger HG, et al. Cloning of a gene highly overexpressed in cancer coding for a novel KH-domain containing protein. Oncogene. 1997;14(22):2729–33.
35. Clauditz TS, Wang CJ, Gontarewicz A, Blessmann M, Tennstedt P, Borgmann K, Tribius S, Sauter G, Dalchow C, Knecht R, et al. Expression of insulin-like growth factor II mRNA-binding protein 3 in squamous cell carcinomas of the head and neck. Journal of oral pathology medicine: official publication of the International Association of Oral Pathologists the American Academy of Oral Pathology. 2013;42(2):125–32.
36. Wang T, Fan L, Watanabe Y, McNeill PD, Moulton GG, Bangur C, Fanger GR, Okada M, Inoue Y, Persing DH, et al. L523S, an RNA-binding protein as a potential therapeutic target for lung cancer. British journal of cancer. 2003;88(6):887–94.

37. Li D, Yan D, Tang H, Zhou C, Fan J, Li S, Wang X, Xia J, Huang F, Qiu G, et al. IMP3 is a novel prognostic marker that correlates with colon cancer progression and pathogenesis. Ann Surg Oncol. 2009;16(12):3499–506.
38. Jeng YM, Chang CC, Hu FC, Chou HY, Kao HL, Wang TH, Hsu HC. RNA-binding protein insulin-like growth factor II mRNA-binding protein 3 expression promotes tumor invasion and predicts early recurrence and poor prognosis in hepatocellular carcinoma. Hepatology. 2008;48(4):1118–27.
39. Suvasini R, Shruti B, Thota B, Shinde SV, Friedmann-Morvinski D, Nawaz Z, Prasanna KV, Thennarasu K, Hegde AS, Arivazhagan A, et al. Insulin growth factor-2 binding protein 3 (IGF2BP3) is a glioblastoma-specific marker that activates phosphatidylinositol 3-kinase/mitogen-activated protein kinase (PI3K/MAPK) pathways by modulating IGF-2. J Biol Chem. 2011;286(29):25882–90.
40. Samanta S, Pursell B, Mercurio AM. IMP3 protein promotes chemoresistance in breast cancer cells by regulating breast cancer resistance protein (ABCG2) expression. J Biol Chem. 2013;288(18):12569–73.
41. Mishima K, Kato Y, Kaneko MK, Nishikawa R, Hirose T, Matsutani M. Increased expression of podoplanin in malignant astrocytic tumors as a novel molecular marker of malignant progression. Acta Neuropathol. 2006;111(5):483–8.
42. Ishii S, Yamashita K, Harada H, Ushiku H, Tanaka T, Nishizawa N, Yokoi K, Washio M, Ema A, Mieno H, et al. The H19-PEG10/IGF2BP3 axis promotes gastric cancer progression in patients with high lymph node ratios. Oncotarget. 2017;8(43):74567–81.
43. Kelemen LE. The role of folate receptor alpha in cancer development, progression and treatment: cause, consequence or innocent bystander? International journal of cancer. 2006;119(2):243–50.
44. Parker N, Turk MJ, Westrick E, Lewis JD, Low PS, Leamon CP. Folate receptor expression in carcinomas and normal tissues determined by a quantitative radioligand binding assay. Anal Biochem. 2005;338(2):284–93.
45. Tomassetti A, Mangiarotti F, Mazzi M, Sforzini S, Miotti S, Galmozzi E, Elwood PC, Canevari S. The variant hepatocyte nuclear factor 1 activates the P1 promoter of the human alpha-folate receptor gene in ovarian carcinoma. Cancer research. 2003;63(3):696–704.
46. Kim M, Pyo S, Kang CH, Lee CO, Lee HK, Choi SU, Park CH. Folate receptor 1 (FOLR1) targeted chimeric antigen receptor (CAR) T cells for the treatment of gastric cancer. PloS one. 2018;13(6):e0198347.
47. Liu J, Wang J, Song Y, Ma B, Luo J, Ni Z, Gao P, Sun J, Zhao J, Chen X, et al. A panel consisting of three novel circulating lncRNAs, is it a predictive tool for gastric cancer? J Cell Mol Med. 2018;22(7):3605–13.

Figures

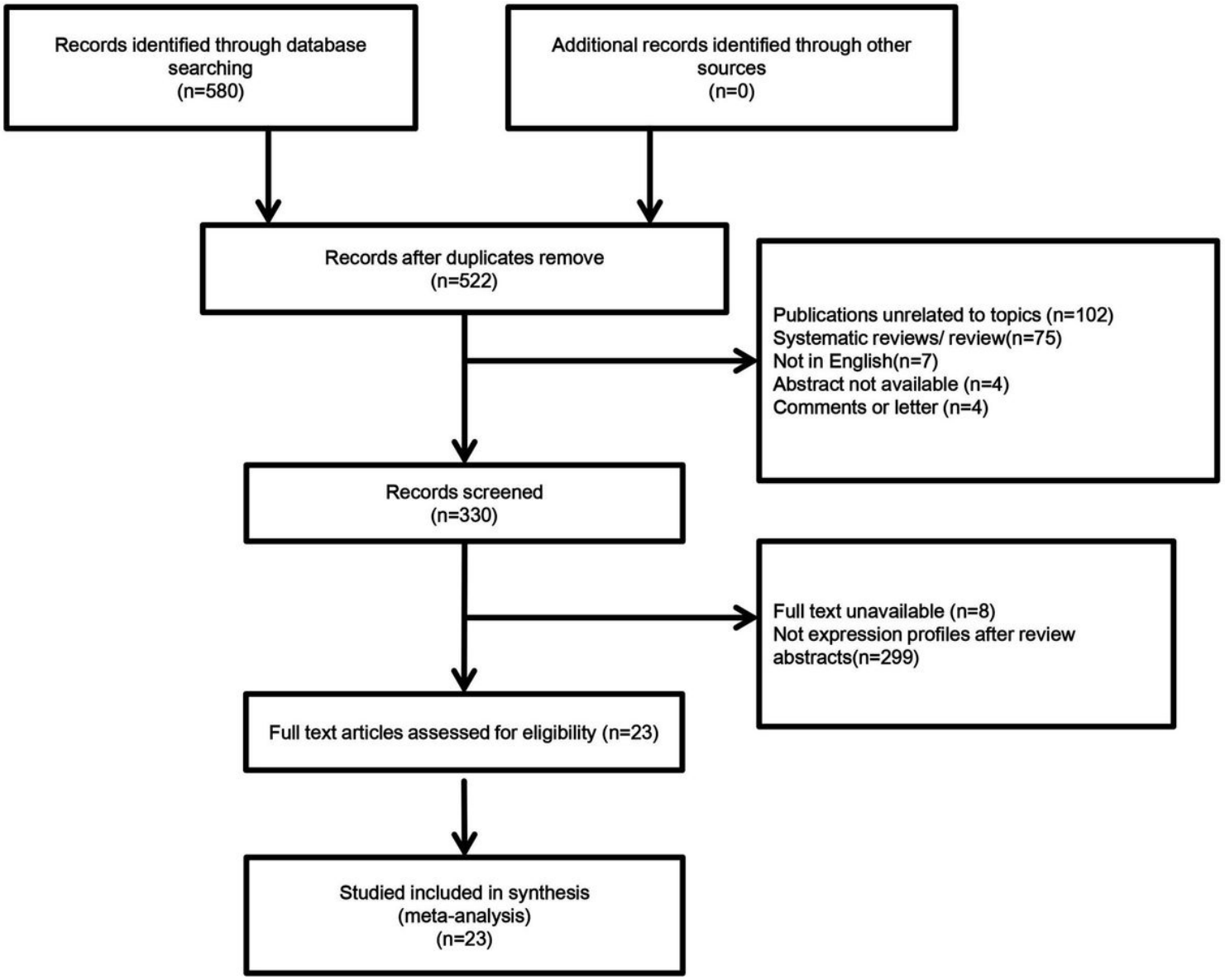

Figure 1

Flowchart showing the procedures for identifying eligible articles for inclusion. This comprehensive algorithm determined at least 23 suitable studies that used LncRNAs as biomarkers for gastric cancer diagnosis

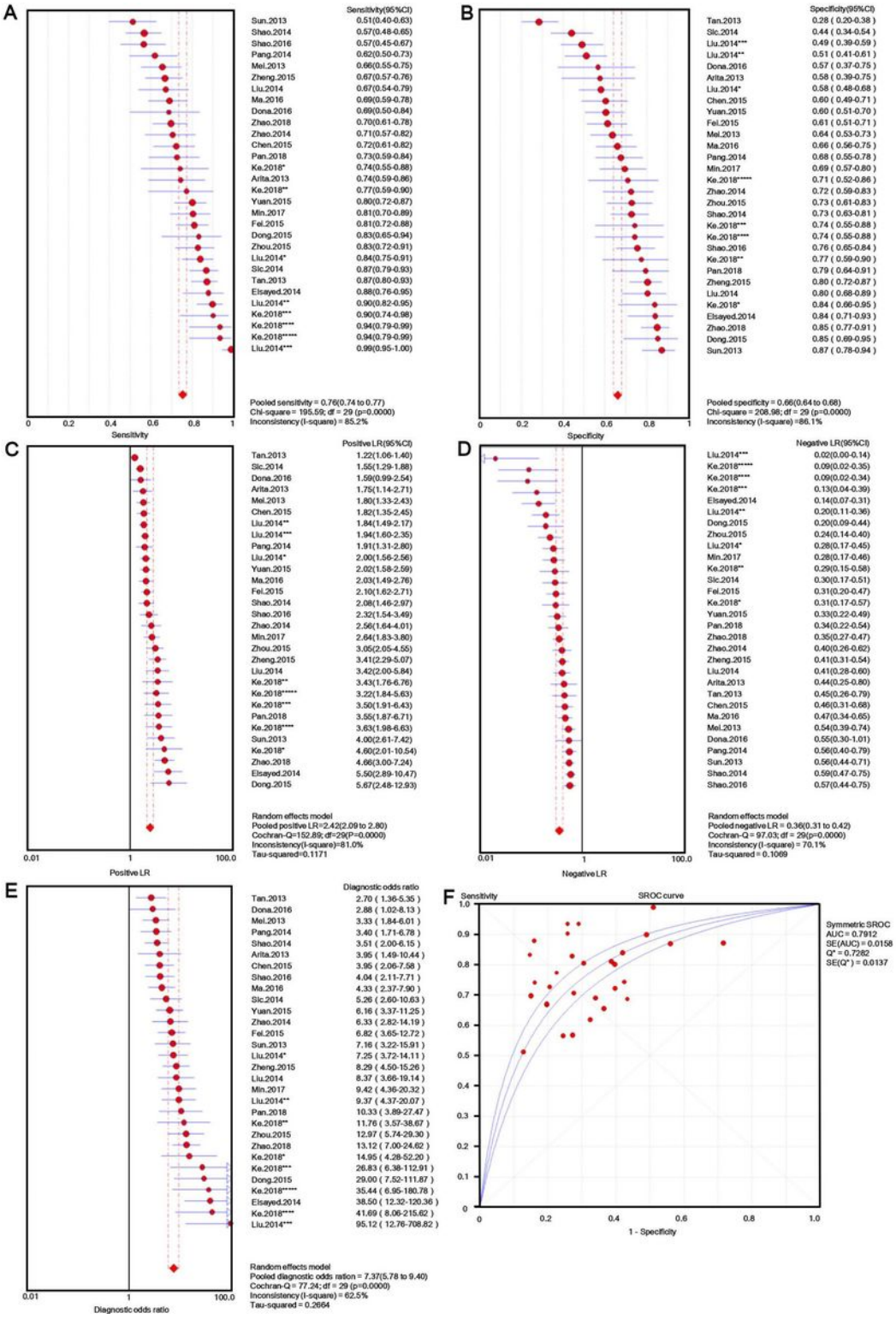

Figure 2

Forest plot of diagnostic value of each index of dysregulated LncRNAs expression in GC. A, Pooled sensitivity of dysregulated LncRNAs expression in diagnosis of GC in all studies. B, Pooled specificity of dysregulated LncRNAs expression in diagnosis of GC in all studies. C, Pooled positive likelihood ratio of dysregulated LncRNAs expression in diagnosis of GC in all studies. D, Pooled negative likelihood ratio of dysregulated LncRNAs expression in diagnosis of GC in all studies. E, Pooled diagnostic odds ratio of

dysregulated LncRNAs expression in diagnosis of GC in all studies. F, sROC of dysregulated LncRNAs expression in diagnosis of GC in all studies.*:Selected LncRNAs coming from one publication

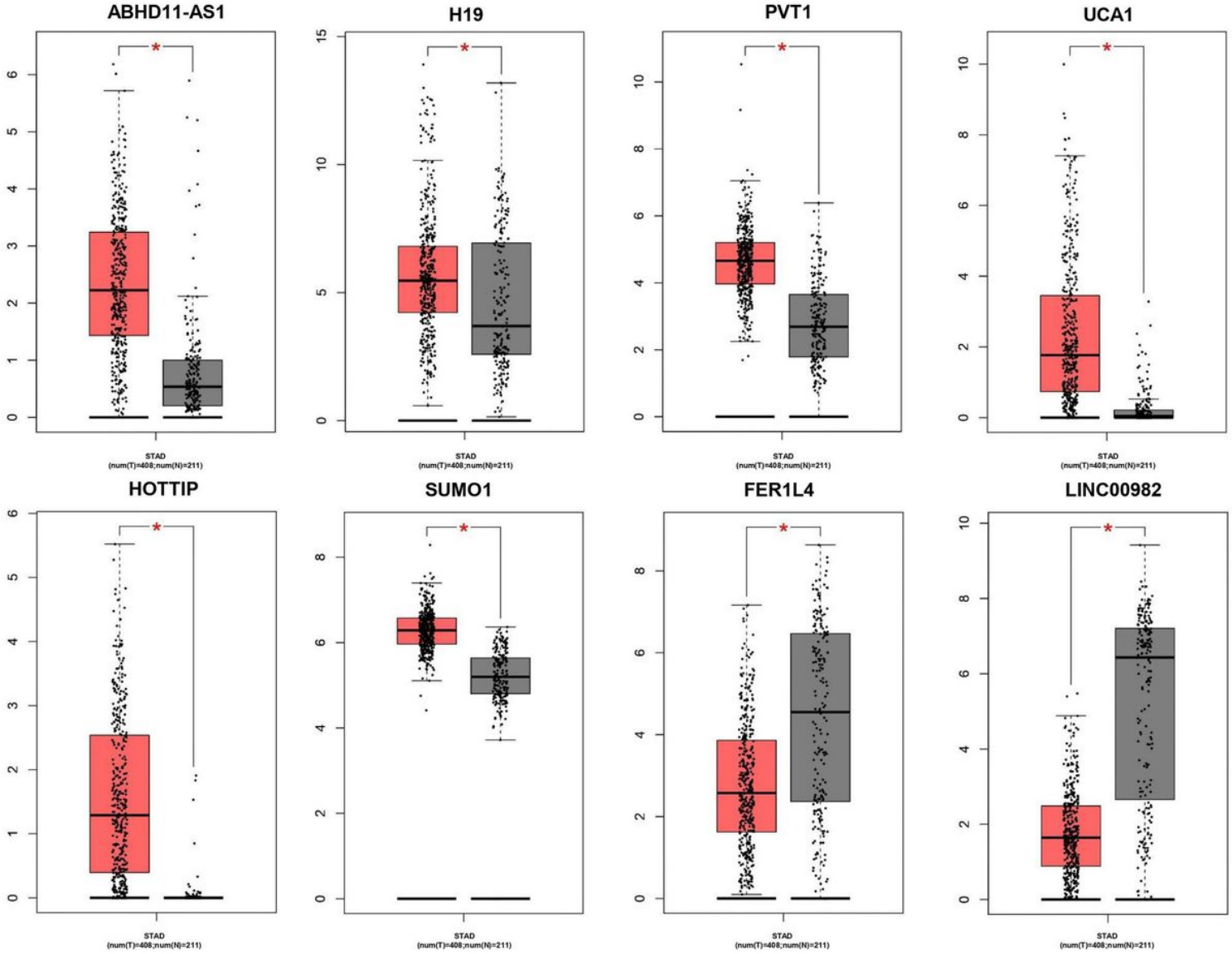

Figure 3

Significantly expressed eight genes in GC specimen compared to normal specimen. To further identify the reported LncRNAs' expression levels for GC and normal people, all genes were further analysed by the GEPIA website. Eight genes were verified in significant expression levels in GC patients compared to healthy people (*P < 0.05, Tumour: Red colour; Normal: Grey colour).

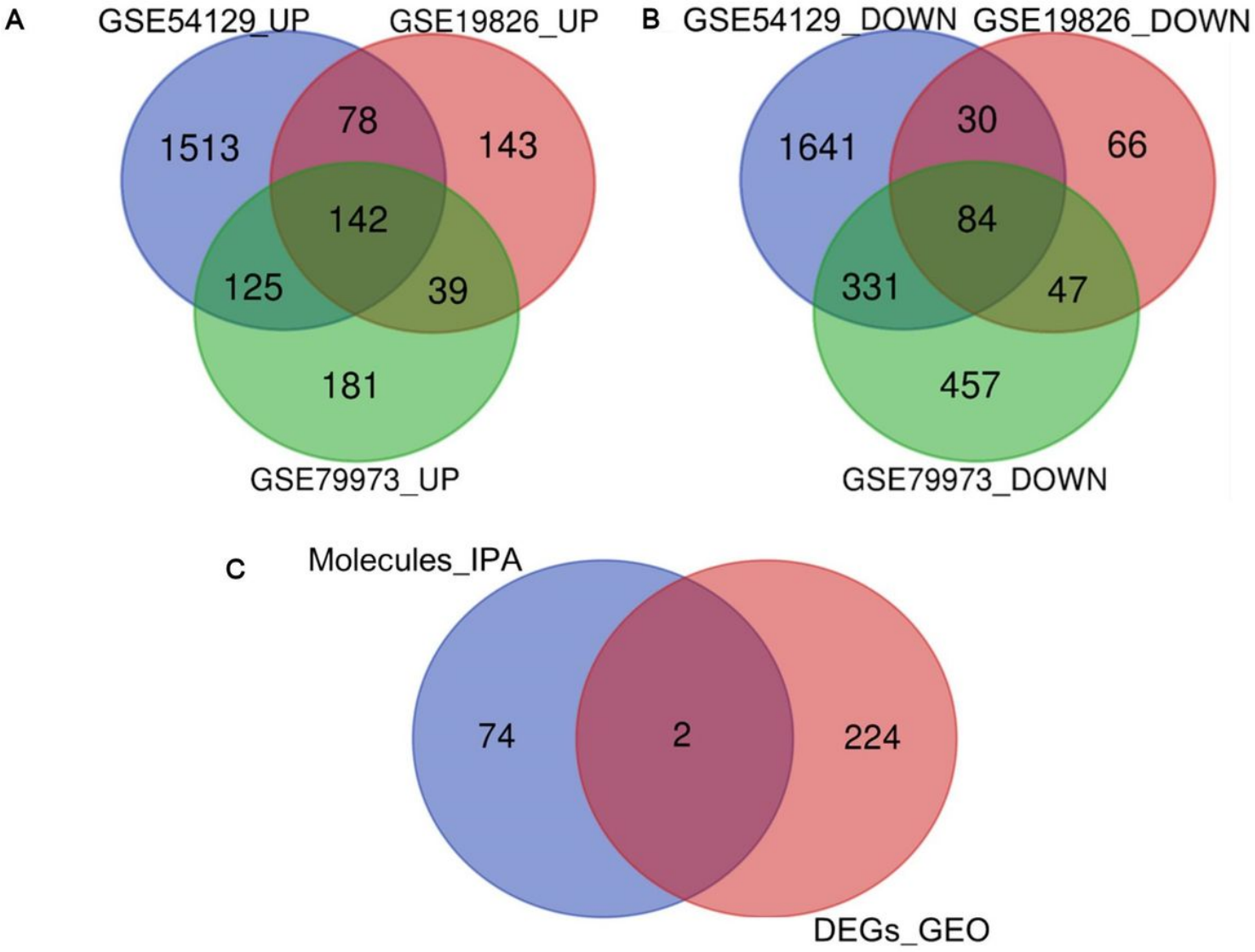

Figure 4

Relationships between DEGs and eight dysregulated LncRNAs in GC patients. A. Venn diagram demonstrating the DEGs upregulated in three datasets (GSE52149, GSE19826 and GSE 79973, $P < 0.05$, LogFC > 2); B. Venn diagram demonstrating the DEGs downregulated in three datasets ($P < 0.05$, LogFC < 2); C. Venn diagram showing the molecules selected from DEGs that interacted with dysregulated LncRNAs in GC patients.

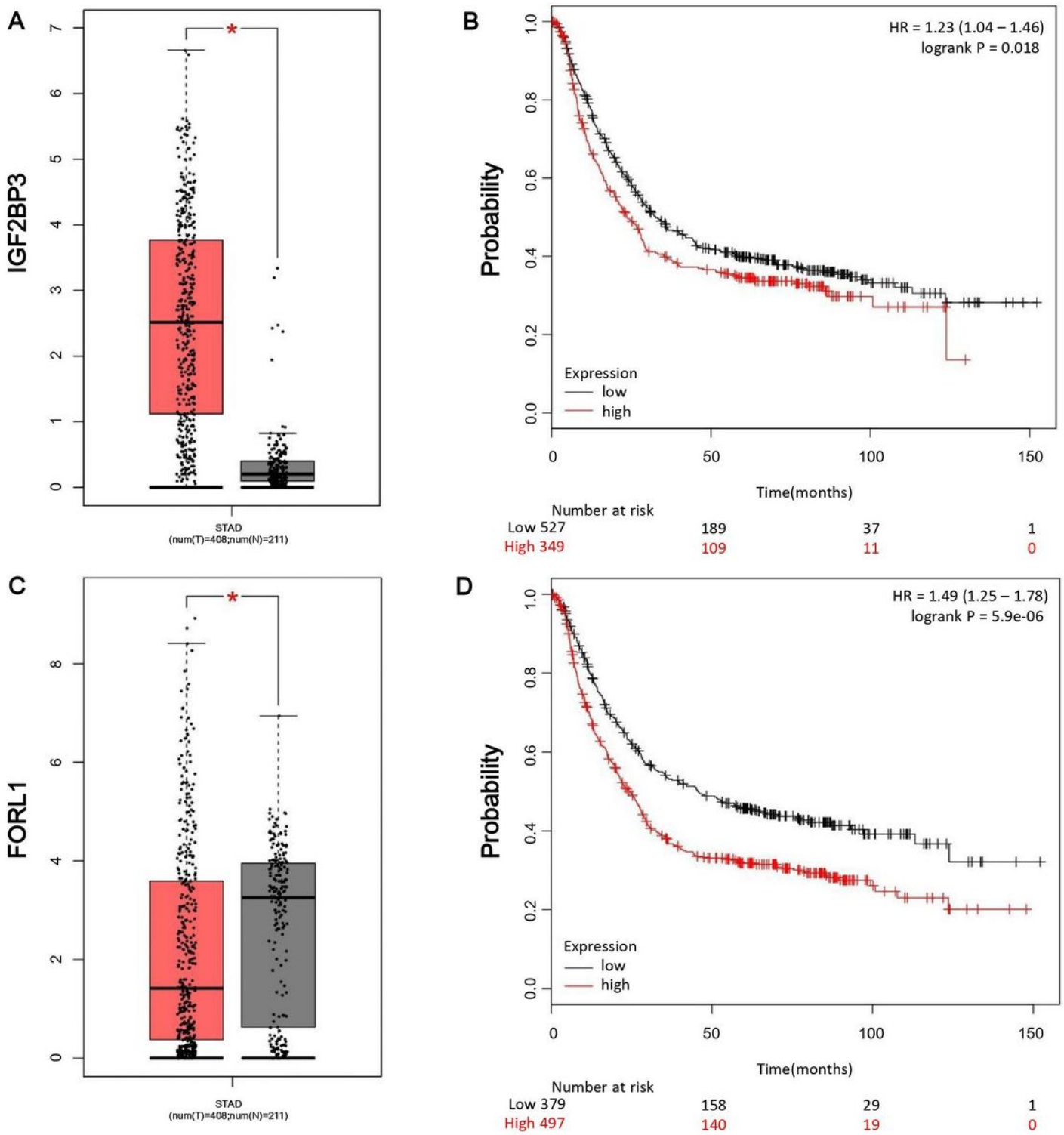

Figure 5

Expression and prognostic value of IGF2BP3 and FOLR1 in GC patients. The expression and potential role of IGF2BP3 and FOLR1 in prognosis in GC were further identified through the GEPIA website and Kaplan Meier Plotter online tools. A. Expression of IGF2BP3 in GC specimen and normal specimen. B. Prognostic relationship of IGF2BP3 in GC patients. C. Expression of FOLR1 in GC specimen and normal specimen. D. Prognostic relationship of FOLR1 in GC patients. (*P < 0.05, Tumour: Red colour; Normal: Grey colour).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile2.jpg
- Additionalfile1.jpg