A Tumor-Specific Prognostic Long Non-Coding RNA Signature in Gastric Cancer

Authors' Contribution:
Study Design A
Data Collection B
Statistical Analysis C
Data Interpretation D
Manuscript Preparation E
Literature Search F
Funds Collection G

BD 1,2   **Wu Ren**
DE 2   **Jian Zhang**
EF 1   **Wei Li**
EG 2   **Zongcheng Li**
EFG 2   **Shuofeng Hu**
A 1   **Jian Suo**
EFG 2   **Xiaomin Ying**

1 Department of Gastrointestinal Surgery, The First Hospital of Jilin University, Changchun, Jilin, P.R. China
2 Beijing Institute of Basic Medical Sciences, Beijing, P.R. China

Background: Aberrant expression of long non-coding RNAs (lncRNAs) is associated with prognosis of gastric cancer, some of which could be further evaluated as potential biomarkers. In this study, we attempted to identify a specific lncRNA signature to predict the prognosis of gastric cancer.

Material/Methods: The genome-wide lncRNA expression in the high-throughput RNA-sequencing data was retrieved from the Cancer Genome Atlas (TCGA). Differential expression of lncRNAs was identified using the Limma package. Survival analysis was conducted by use of univariate and multivariate Cox regression models. Functional enrichment analysis of lncRNAs was based on co-expressed mRNAs. DAVID was used to perform gene ontology and KEGG pathway analysis.

Results: A total of 452 differentially expressed lncRNAs between gastric cancer and matched normal tissues were screened, of which 76 lncRNAs were identified to be gastric cancer-specific from a pan-cancer analysis of 12 types of human cancer. Among these 76 gastric cancer-specific lncRNAs, 5 lncRNAs (CTD-2616J11.14, RP1-90G24.10, RP11-150O12.3, RP11-1149O23.2, and MLK7-AS1) were significantly associated with the overall survival of patients with gastric cancer. A gastric cancer-specific 5-lncRNA signature was deduced to divide the patients into high- and low-risk groups with significantly different survival times ($P<0.0001$). Multivariate Cox regression analysis showed that this 5-lncRNA signature was an independent predictor of prognosis. Functional enrichment analysis of the 5 lncRNAs showed that they were mainly involved in DNA replication, mitotic cell cycle, programmed cell death, and RNA splicing.

Conclusions: Our results suggest that this tumor-specific lncRNA signature may be clinically useful in the prediction of gastric cancer prognosis.

MeSH Keywords: **Biological Markers • Prognosis • RNA, Long Noncoding • Stomach Neoplasms**

Full-text PDF: http://www.medscimonit.com/abstract/index/idArt/901190

  2786   3   4   41

Background

Gastric cancer significantly contributes to cancer-related deaths worldwide; in 2012, there were more than 950 000 new cases and 720 000 deaths occurring globally [1]. Gastric cancer is frequently under-diagnosed clinically, because disease-related symptoms often appear late or are atypical during gastric cancer development, leading to most cases being diagnosed at an advanced stage of the disease, thereby narrowing options for therapy and worsening prognosis [2,3]. At present, surgery is the best method of cure; however, patients with advanced-stage gastric cancer may not be surgically cured, and chemo- or radiotherapy in such patients often leads to resistance and disease relapse [4]. Although recent advances in the understanding of biological properties and treatment strategies have helped to reduce the incidence and mortality rates of gastric cancer, the overall 5-year survival rate is still only approximately 25% [5,6]. Therefore, there is an urgent need to identify and evaluate molecular biomarkers for prediction of prognosis and treatment response in patients with gastric cancer.

In the last few years, along with an extensive characterization of the protein-coding genome in gastric cancer, increasing attention has been focused on research on lncRNAs, which are a class of non-coding RNAs with a length of more than 200 nucleotides. LncRNAs play a regulatory role at transcriptional, post-transcriptional, and epigenetic levels to modify the expression of protein-coding genes [7], and they mediate various biological processes, such as cell proliferation, differentiation, signal transduction, and apoptosis [8–10]. Recent studies demonstrated the significance of aberrant lncRNA expression in human cancers [11] and as oncogenes or tumor suppressors in cancer development and progression [12]. Several studies have assessed the aberrant lncRNA expression in gastric cancer and explored their clinical relevance in diagnosis, prediction of prognosis, and treatment efficacy [13–17]. Thus, a tissue-specific lncRNA alteration could make them ideal biomarkers for gastric cancer [18].

In this study, we aimed to establish a tumor-specific lncRNA prognostic signature for patients with gastric cancer. We used RNA-sequencing data and clinical information of patients with gastric cancer from TCGA. Our study design comprised: i) identifying differentially expressed lncRNAs between gastric cancer and matched normal tissue; ii) comparing differentially expressed lncRNAs with that of other 11 cancer types to identify gastric cancer-specific lncRNAs; and iii) screening and evaluating prognosis-associated lncRNAs as a combined signature from gastric cancer-specific lncRNAs.

Material and Methods

Data collection

LncRNA expression profiles of tumor and normal samples from patients in TCGA were retrieved from the Atlas of Non-coding RNAs in Cancer (TANRIC, *http://bioinformatics.mdanderson.org/main/TANRIC:Overview*) database [19]. The database includes 12 727 lncRNAs identified within the TCGA RNA-sequencing datasets. The TANRIC annotations relied on human lncRNAs from the GENCODE Resource (version 19), and any lncRNAs that overlapped with any given mRNAs were filtered out. Reads per kilobase per million mapped reads (RPKM) values were calculated using TCGA RNA-sequencing data in the BAM files. The data also included an additional 11 TCGA solid tumor types: thyroid carcinoma (THCA), prostate adenocarcinoma (PRAD), kidney chromophobe (KICH), invasive breast carcinoma (BRCA), kidney renal papillary cell carcinoma (KIRP), head and neck squamous cell carcinoma (HNSC), hepatocellular carcinoma (LIHC), kidney renal clear cell carcinoma (KIRC), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), and bladder urothelial carcinoma (BLCA), each of which has more than 10 normal tissue samples available in the database. Profiling of mRNA expression and clinical data in TCGA data set were directly derived from the UCSC Cancer Genomics Browser [20].

Patients and lncRNA data processing

Patients meeting the following criteria were included in the study: i) a histological diagnosis of primary gastric cancer; ii) having no history of neoadjuvant therapy; iii) lncRNA expression profile of tumor sample; and iv) available follow-up data. A total of 254 patients with gastric adenocarcinoma (cohort T) were included in the study, with corresponding clinical data including sex, age, tumor location, American Joint Committee on Cancer (AJCC) pathologic stage, invasion depth, lymph node metastasis, distal metastasis, differentiation grade, microsatellite instability (MSI) status, viral status, and relapse status (Table 1). Among these 254 patients, matched normal gastric mucosae were from 29 patients (cohort N). Another 11 cancer types with more than 10 normal samples were included in the pan-cancer analysis, and patients having tumors belonging to any of the 11 cancer types were selected following the criteria: i) a histological diagnosis of primary cancer and ii) available lncRNA expression profiles of tumor samples. Matched normal samples for these 11 cancer types were from these selected patients. The profiles of lncRNA and mRNA expression were normalized as RPKM values. Since many lncRNAs had a low expression level in certain tissues, lncRNA levels with a mean RPKM value less than 0.3 across all samples were excluded from the data analysis.

Table 1. Clinical characteristics of patients with gastric cancer.

Category	Cohort T (n=254) n (%)		Cohort N (n=29) n (%)		P-value*
Gender					0.28
Male	158	(62.20)	21	(72.41)	
Female	96	(37.80)	8	(27.59)	
Age (years)					0.66
≥65	138	(54.33)	17	(58.62)	
<65	116	(45.67)	12	(41.38)	
Median age	66		67		
Range	34–90		46–81		
Tumor location					0.006
Gastroesophageal junction	12	(4.72)	6	(20.69)	
Cardia/proximal	38	(14.96)	3	(10.34)	
Fundus/body	97	(38.19)	13	(44.83)	
Antrum/distal	100	(39.37)	7	(24.14)	
Unknown	7	(2.76)	0		
AJCC pathologic stage					0.14
I	35	(13.78)	3	(10.34)	
II	96	(37.79)	17	(58.62)	
III	99	(38.98)	6	(20.69)	
IV	21	(8.27)	3	(10.34)	
Unknown	3	(1.18)	0		
Invasion depth					0.37
T1	10	(3.94)	1	(3.45)	
T2	65	(25.59)	10	(34.48)	
T3	105	(41.34)	14	(48.28)	
T4	72	(28.34)	4	(13.79)	
Unknown	2	(0.79)	0		
Lymph node metastasis					0.013
N0	87	(34.25)	8	(27.59)	
N1	69	(27.16)	11	(37.93)	
N2	45	(17.72)	10	(34.48)	
N3	50	(19.69)	0		
Unknown	3	(1.18)	0		

Table 1 continued. Clinical characteristics of patients with gastric cancer.

Category	Cohort T (n=254) n (%)		Cohort N (n=29) n (%)		P-value*
Distal metastasis					0.56
M0	228	(89.76)	26	(89.65)	
M1	16	(6.30)	1	(3.45)	
Unknown	10	(3.94)	2	(6.90)	
Tumor differentiation					0.40
G1	5	(1.97)	0		
G2	76	(29.92)	12	(41.38)	
G3	168	(66.14)	17	(58.62)	
Unknown	5	(1.97)	0		
MSI status					0.46
MSI-H	48	(18.90)	6	(20.69)	
MSI-L	40	(15.75)	7	(24.14)	
MSS	165	(64.96)	16	(55.17)	
Unknown	1	(0.39)	0		
Vital status					0.73
Death	69	(27.17)	7	(24.14)	
Live	185	(72.83)	22	(75.86)	
Relapse status					0.065
Relapse	42	(16.53)	1	(3.45)	
Not relapse	175	(68.90)	23	(79.31)	
Unknown	37	(14.57)	5	(17.24)	

*P-value was calculated using Chi-square test.

Statistical analysis

The differences of clinical variables (sex, age, tumor location, AJCC pathologic stage, invasion depth, lymph node metastasis, distal metastasis, MSI status, differentiation grade, viral status, and relapse status) between cohort T and cohort N were analyzed using the chi-square test. Differential lncRNA expression between tumor and normal tissues of the 12 cancer types was analyzed using the R (version 3.2.2)/Bioconductor software package Limma [21] with the criteria of adjusted P-value <0.01 and fold change >2, respectively. The Limma uses linear models and empirical Bayes paired moderated t-statistics and F-statistics. The unsupervised hierarchical cluster analysis was carried out using heatmap.2 function of the R/package gplots with complete linkage.

The association of lncRNA expression with the overall survival of patients was analyzed using the univariate Cox proportional hazards regression model. After selecting lncRNAs with a criteria of P-value <0.05, multivariate Cox proportional hazards regression model was used to calculate a relative regression coefficient (ri) for each lncRNA and to combine lncRNA expression (Exp [i]) into a linear risk score model, risk score = $\sum_{i=1}^{n} \mathrm{Exp(i)} * \mathrm{ri}$ [22]. The Kaplan-Meier method and log-rank test were performed to analyze the survival of patients. Clinical variables were included in a univariate Cox regression model to assess their association with overall survival. A multivariate Cox regression model of the risk score with clinical variables was used to assess the independence of the lncRNA risk score. Survival analysis was performed using SPSS software (version 19) and R/package survival.

Functional annotation of lncRNAs was based on co-expressed mRNAs using a 'guilt-by-association' strategy [23]. LncRNA-mRNA pairs with the top 1% absolute Spearman's correlation coefficient were filtered out. Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of these lncRNA co-expressed mRNAs were performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID, *http://www.david.niaid.nih.gov*) (version 6.8) [24]. The enriched results were restricted to GO biological process and KEGG pathway terms, and terms were clustered using the "the functional annotation clustering" tool. The GO biological process and KEGG pathway terms with adjusted *P*-value <0.05 were considered to be significant. In this study, *P*-value was adjusted using the Benjamini-Hochberg method.

Results

Patient characteristics

All 254 gastric cancer patients in this study were histologically diagnosed with gastric adenocarcinoma, without a history of neoadjuvant therapy. The mean ±SD of the overall survival time was 13.47±12.06 months and the mean ±SD of the relapse-free survival time was 12.65±9.97 months. Among these 254 patients (cohort T), matched normal tissues from 29 patients (cohort N) were used for differential lncRNA expression analysis between gastric cancer tissues and normal tissues. As shown in Table 1, there was no significant difference between cohort T and cohort N in sex (*P*=0.28), age (*P*=0.66), AJCC pathologic stage (*P*=0.14), invasion depth (*P*=0.37), distal metastasis (*P*=0.56), MSI status (*P*=0.46), differentiation grade (*P*=0.40), viral status (*P*=0.73), and relapse status (*P*=0.065).

Differentially expressed lncRNAs between gastric cancer tissues and normal tissues

We identified 452 differentially expressed lncRNAs between gastric cancer tissues and normal tissues based on the expression profile of lncRNAs with the criteria of fold change >2 and adjusted *P*-value <0.01. As shown in Figure 1A, unsupervised hierarchical clustering clearly divided the tumor and normal samples with the expression of the altered lncRNAs. A total of 100 lncRNAs showed more than 8-fold expression change; the top 5 upregulated lncRNAs were HOTAIR, HOXC-AS1, RP1-170O19.14, RP11-400N13.3, and AC012363.4, while the top 5 downregulated lncRNAs were AC053503.6, PGM5-AS1, CTC-276P9.1, CTD-2231H16.1, and SOX21-AS1 (Figure 1B).

Pan-cancer analysis of differentially expressed lncRNAs

We then assessed the tissue-specific alteration pattern of these 452 aberrant lncRNAs in gastric cancer *vs.* the other 11

types of cancer and their matched normal tissues. A total of 4006 tumor samples and 507 partially matched normal samples from the additional 11 cancer types were included in this pan-cancer analysis. We found that 30 lncRNAs were differentially expressed in more than 7 tumor types, 123 lncRNAs in 5 to 7 tumor types, 223 lncRNAs in 2 to 4 tumor types, and 76 lncRNAs only in gastric cancer (Figure 1C). The 30 lncRNAs that were differentially expressed in more than 7 tumor types may have an oncogenic or tumor suppressor role in these human cancers; for example, PVT1 and FAM83H-AS1 have been reported to be biomarkers in human cancers [18,25]. The 76 lncRNAs that were differentially expressed only in gastric cancer could be potential gastric cancer-specific biomarkers, although their functions remain unknown.

Establishment of lncRNA signature associated with overall survival of patients with gastric cancer

These 76 gastric cancer-specific lncRNAs were further assessed for their association with the overall survival of patients with gastric cancer. Five lncRNAs (CTD-2616J11.14, RP1-90G24.10, RP11-150O12.3, RP11-1149O23.2, and MLK7-AS1) showed a significant *P*-value of <0.05 (Table 2). Then, a relative regression coefficient was calculated by multivariate Cox regression analysis for each lncRNA, and a linear model was generated combining the 5 lncRNAs. A risk score was calculated for each patient in the TCGA cohort using this model, and patients could be divided into low- (n=127) and high-risk groups (n=127) using the median of risk scores as the cut-off value (Figure 2A). Patients in the high-risk group had a poor median survival time compared to those in the low-risk group (*P*<0.0001 using the log-rank test; Figure 2B). Recurrence data were also available for 217 patients, and the lncRNA risk score was effective in dividing patients into high- and low-risk groups using the median as the cut-off value (*P*=0.021 using the log-rank test; Figure 2C).

Association of clinical features with overall survival and validation of 5-lncRNA signature as an independent predictor

Univariate Cox regression analysis was performed to evaluate the association between clinical features and overall survival. We found that AJCC pathologic stage, invasion depth, lymph node metastasis, and differentiation grade were statistically significant for the prediction of overall survival of patients (*P*<0.05; Table 3 and Figure 3), whereas age, sex, tumor location, and MSI status were not. However, the number of patients with distal metastasis was small, and the association of distal metastasis with overall survival was not assessed.

We further performed multivariate Cox regression analysis incorporating the 5-lncRNA risk score with clinical features that
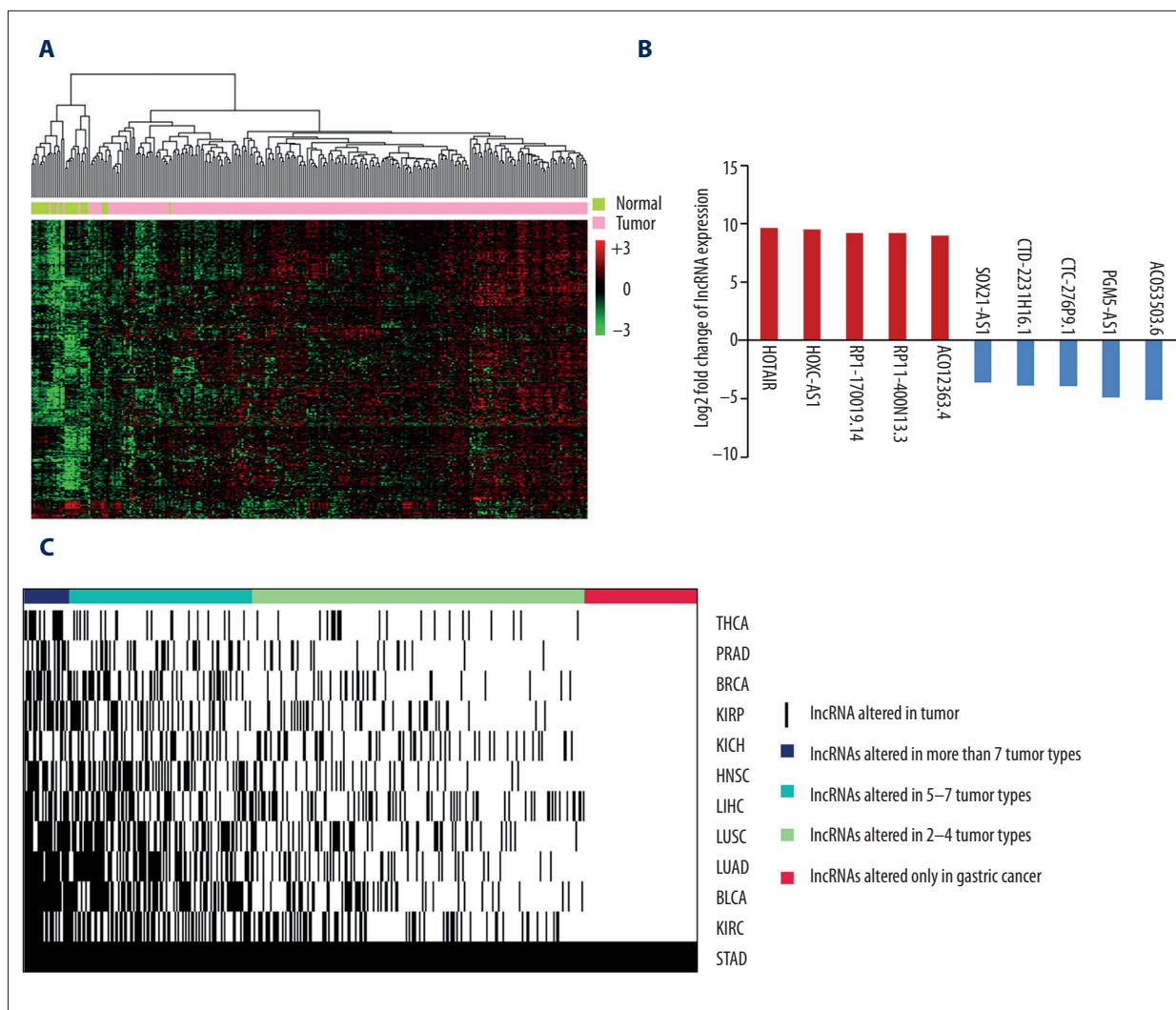
Figure 1. Differentially expressed lncRNAs in gastric cancer vs normal tissues. (**A**) Unsupervised hierarchical cluster analysis of the differentially expressed lncRNAs between gastric cancer and normal tissues. (**B**) The top 5 upregulated lncRNAs (red) and the top 5 downregulated lncRNAs (blue). (**C**) Pan-cancer analysis of differentially expressed lncRNAs in gastric cancer and the other 11 cancer types. Black bar represents differentially expressed lncRNAs in a specific cancer type.

Table 2. Association of LncRNA expression with overall survival of gastric cancer patients.

Ensembl ID	LncRNA	Chromosome location	Univariate Cox regression analysis				Relative coefficient
			HR	95% CI of HR	Coefficient	P-value	
ENSG00000268889	CTD-2616J11.14	19: 51,897,742-51,906,904	0.325	0.128–0.827	−1.12	0.018	−0.89
ENSG00000242082	RP1-90G24.10	22: 32,601,102-32,665,653	0.299	0.109–0.818	−1.21	0.019	−0.82
ENSG00000254290	RP11-150O12.3	8: 37,454,998-37,457,376	0.752	0.589–0.958	−0.29	0.021	−0.22
ENSG00000253930	RP11-1149O23.2	8: 23,046,792-23,048,188	0.736	0.546–0.993	−0.31	0.045	−0.24
ENSG00000238133	MLK7-AS1	2: 174,031,174-174,146,764	0.696	0.498–0.971	−0.36	0.033	−0.29

HR – hazard ratio; CI – confidential interval.
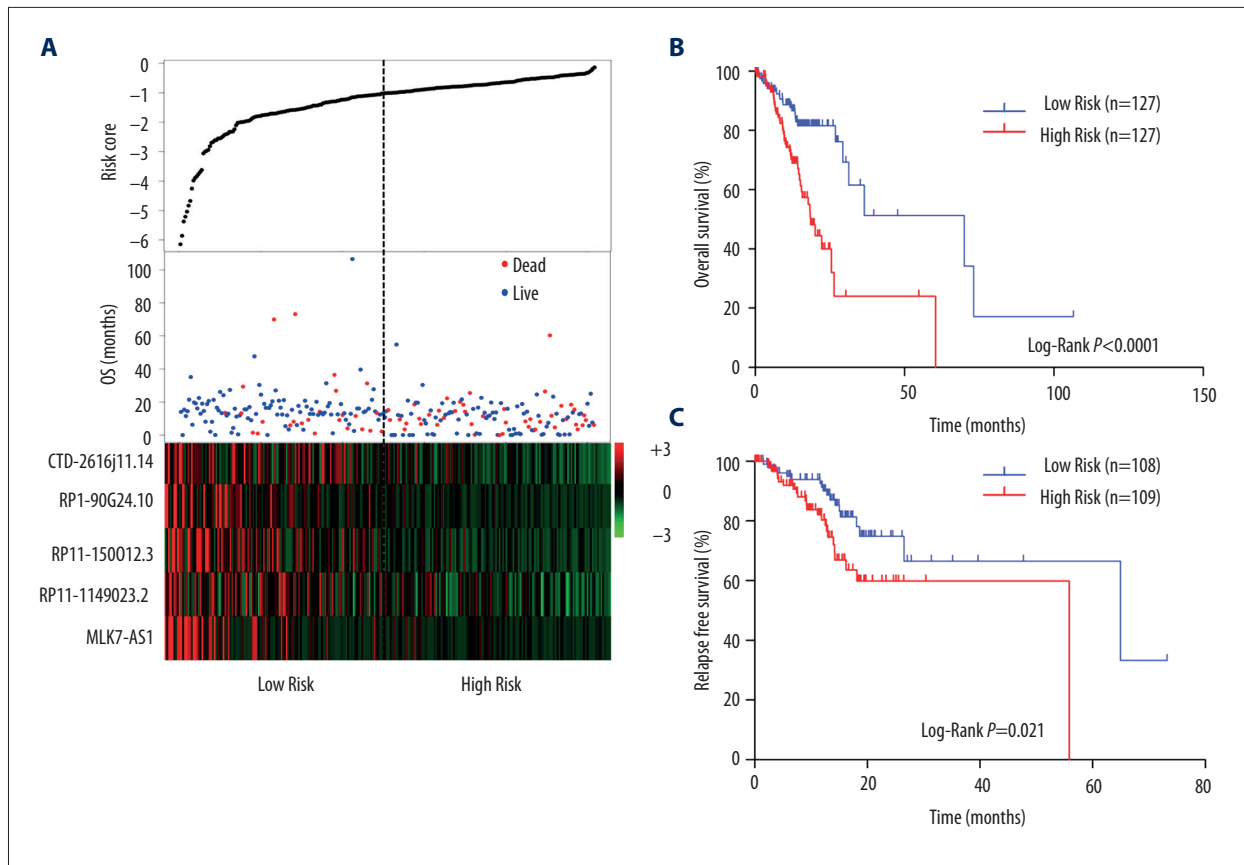
Figure 2. Prognostic value of the 5-lncRNA signature in gastric cancer patients. (**A**) Risk score distribution (top), overall patient survival (middle), and the expression heatmap of the 5 lncRNAs (bottom). (**B**) Kaplan-Meier survival curve analysis of overall survivals. Patients with gastric cancer can be divided into high-risk (red) and low-risk (blue) groups by the 5-lncRNA signature. (**C**) Kaplan-Meier survival curve analysis of relapse-free survival. Patients with gastric cancer can be divided into high-risk (red) and low-risk (blue) groups. *P*-value was calculated using the log-rank test.

were significantly associated with overall survival, to assess whether this 5-lncRNA signature was an independent prognostic factor. The result showed that this 5-lncRNA risk score (Hazard Ratio [HR]=2.34, *P*=0.002) was an independent prognostic predictor for the overall survival of patients with gastric cancer (Table 3).

Functional annotation of the five-lncRNA signature co-expressed genes

Next, the 'guilt-by-association' strategy was used to predict the functions of lncRNAs. For these 5 prognosis-associated lncRNAs, the Spearman's correlation coefficients with each protein-coding gene were calculated and functional annotation was completed according to those significant genes. Functional enrichment analysis showed that co-expressed genes of the 5 lncRNAs were enriched in 30 GO biological process terms and 2 KEGG pathway terms (adjusted *P*-value <0.05), most of which were grouped into 3 functional clusters of DNA replication, mitotic cell cycle and programmed cell death, and RNA splicing

(Figure 4). In addition, RP11-1149O23.2 was localized at the minus strand relative to a protein-coding gene *TNFRSF10A*, the protein of which is a pro-apoptosis receptor to mediate the extrinsic apoptosis pathway [26]. RP11-1149O23.2 has a high co-expression coefficient of 0.81 with TNFRSF10A, suggesting the possibility of RP11-1149O23.2 in *cis*-regulation of TNFRSF10A. As is well-known, the apoptotic process is an important format of programmed cell death [27]; suggesting a potential role of RP11-1149O23.2 in the process of programmed cell death.

Discussion

In recent decades, considerable efforts have been made to improve the clinical outcome of gastric cancer; however, the overall survival of patients with gastric cancer is still poor [28]. In daily clinical practice, pathologic stage is still commonly used to index prognosis estimation and guide treatment for gastric cancer; however, tumor heterogeneity frequently occurs in patients within similar tumor stages [29]. Thus, reliable prognostic

**Table 3.** Univariate and multivariate Cox regression analysis.

Variables	Univariate analysis			Multivariate analysis		
	HR	95% CI of HR	*P*-value	HR	95% CI of HR	*P*-value
Age: ≥65 *vs.* <65	1.08	0.662–1.752	0.76			
Gender: Male *vs.* Female	0.96	0.583–1.581	0.87			
MSI: MSI-L+MSS *vs.* MSI-H	1.34	0.720–2.522	0.35			
Pathologic Stage: III + IV *vs.* I+II	2.20	1.328–3.639	0.002	1.46	0.722–2.962	0.29
Grade: Grade 3 *vs.* Grade 1/2	1.81	1.041–3.148	0.036	1.42	0.775–2.627	0.25
Invasion: T3–4 *vs.* T1–2	1.89	1.070–3.355	0.028	1.20	0.632–2.283	0.58
Lymph node: N1–3 *vs.* N0	1.97	1.108–3.516	0.021	1.14	0.531–2.486	0.72
Location: Distal *vs.* Proximal	0.85	0.487–1.504	0.58			
LncRNA risk score: High *vs.* Low	2.96	1.779–4.938	3.19E–05	2.34	1.361–4.036	0.002

HR – hazard ratio; CI – confidence interval.
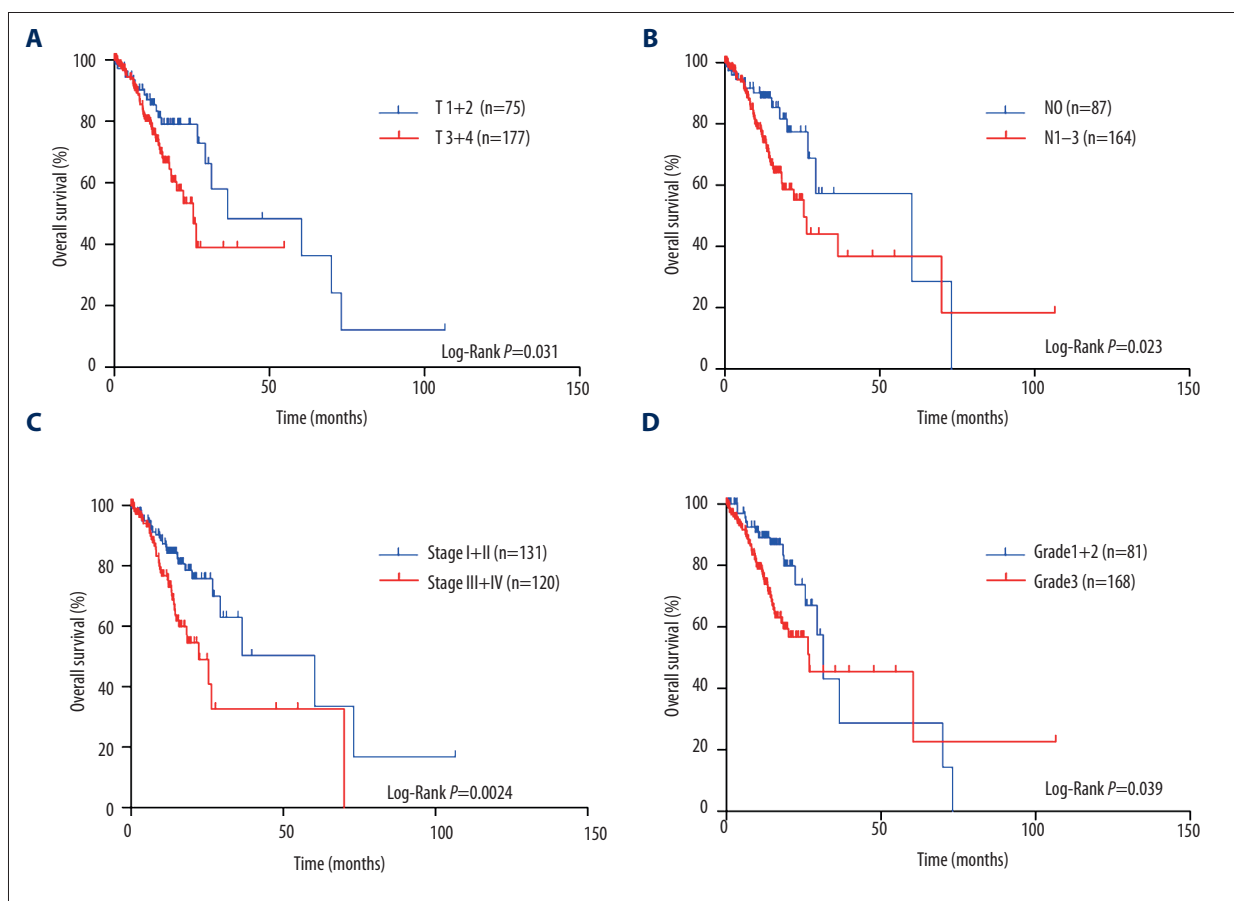
Figure 3. Kaplan-Meier survival curves stratified by clinical covariates for patients with gastric cancer. (**A**) Invasion depth. T1+T2 *vs.* T3+T4. (**B**) Lymph node metastasis. N0 *vs.* N1–3. (**C**) AJCC pathologic stage. Stage I + Stage II *vs.* Stage III + Stage IV. (**D**) Tumor differentiation: grade 1 + grade 2 vs grade 3. *P*-value was calculated using the log-rank test.
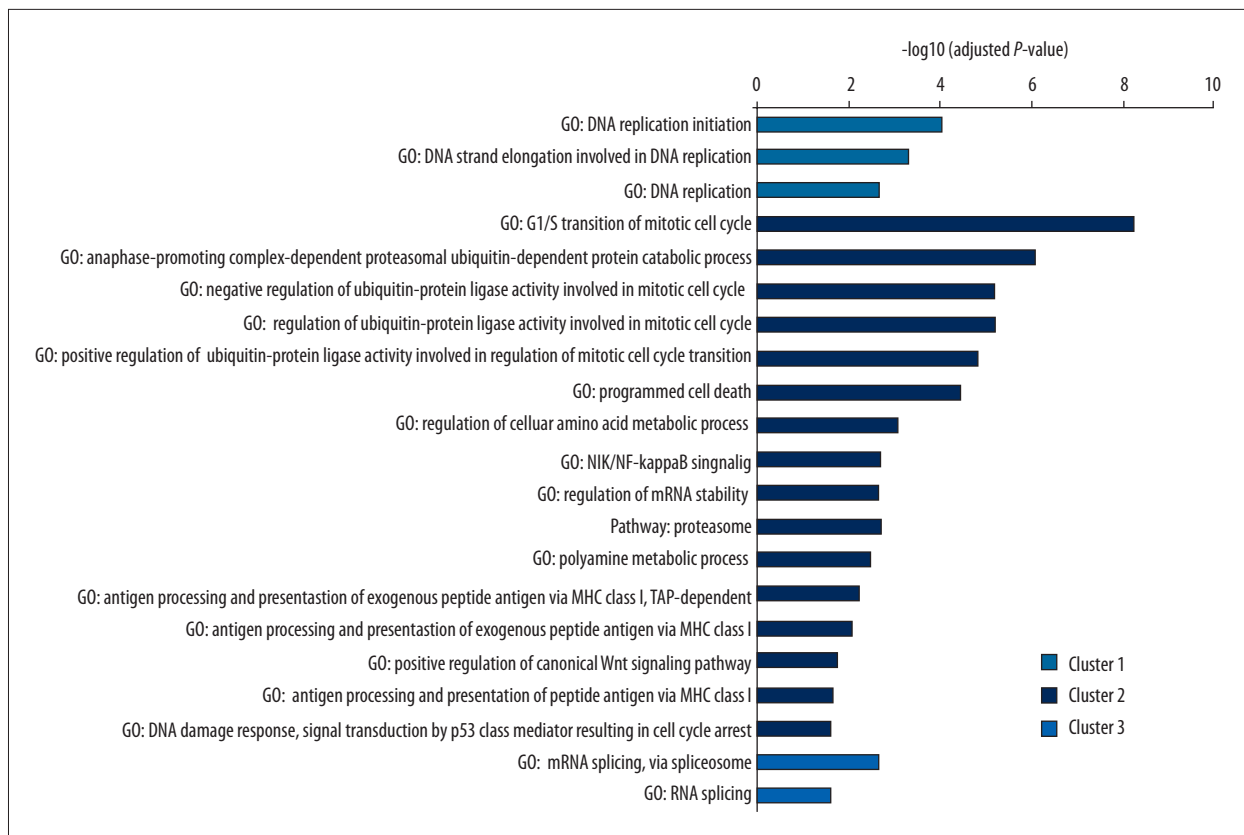
Figure 4. Functional enrichment analysis of the 5-lncRNA signature. Significant gene ontology biological process and KEGG pathway terms were grouped into 3 clusters, which were mainly involved in the regulation of DNA replication (cluster 1), mitotic cell cycle and programmed cell death (cluster 2), and RNA splicing (cluster 3), respectively.

biomarkers are required. In this study, a 5-lncRNA prognostic signature was identified after differential expression, as well as pan-cancer and survival analyses, which was then confirmed to be an independent prognostic predictor for patients with gastric cancer. This study explored the potential of combined lncRNA signature to predict the prognosis of gastric cancer.

Considering the importance of lncRNAs in cancer development and progression, the altered lncRNA expression may be an indicator of the intrinsic characteristics of tumor cells. Altered lncRNA exists in solid cancer tissue, digestive juice, plasma, and urine, which is linked to the occurrence, progression, and outcome of cancer [17,30]. To date, most cancer studies of lncRNA were designed for a general point of view and the tissue-specific alteration pattern of lncRNAs was ignored, resulting in no assurance of finding more putative biomarkers. To the best of our knowledge, ours is the first study to analyze differentially expressed lncRNAs in 12 cancer types, including gastric cancer, to distinguish gastric cancer-specific lncRNAs, and a total of 76 lncRNAs were found to be only altered in gastric cancer. Thus, further study of these 76 lncRNAs could provide more information about their possible use as putative biomarkers for gastric cancer.

Previous studies have identified the association between combined lncRNA expression and the prognosis of colorectal cancer [31], glioblastoma [32], breast cancer [33], lung cancer [34], and esophageal cancer [35]. In gastric cancer, aberrant lncRNAs have been shown to correlate with overall survival or treatment response. For example, Hu et al. used microarray and *in situ* hybridization analysis, and found that GAPLINC was highly expressed in gastric cancer and defines a group of patients with poor prognosis [31]. Zhang et al. highlighted the role of TUG1 in regulating cell cycle and its overexpression in accordance with poor survival [16]. GAS5 is decreased in gastric cancer, which is associated with poor prognosis [36]. Similar to gene or miRNA expression signature for gastric cancer, a combined lncRNA signature may substantially improve the prediction of clinical outcome [37,38]. In the present study, a gastric cancer-specific prognostic risk score model was constructed using 5 differentially expressed lncRNAs and the TCGA cohort was divided into low- and high-risk groups. We found that the 5-lncRNA risk score was independent of pathologic stage and tumor differentiation. The HR of the signature was markedly higher than that of AJCC pathologic stage (2.96 *vs.* 2.20). Owing to the specificity of the 5 lncRNAs, the signature may be more feasible to use in clinical practice. In addition, the risk score

also showed its usefulness in predicting relapse-free survival, indicating its value in assessing treatment efficacy.

The carcinogenesis of gastric cancer is a multi-step process caused by changes in the genetics and epigenetics [39]. However, as the functions of most lncRNAs remain unknown, computational algorithm analysis is an appropriate way to predict and provide estimations for lncRNA function [40]. In our study, we utilized this 5-lncRNA signature as an example to explore their mechanisms using co-expression analysis of lncRNAs and mRNAs. Functional enrichment analysis of lncRNA co-expressed mRNAs revealed that this 5-lncRNA signature is mainly involved in DNA replication, mitotic cell cycle, programmed cell death, and RNA splicing. All these biological processes play essential roles in the pathological progression of gastric cancer [41]. Further studies are warranted to confirm the functions of the 5 lncRNAs and to explore the underlying mechanism in gastric cancer.

Our study is just proof-of-principle and has some limitations. For example, the follow-up period was relatively short in the TCGA dataset and the censoring rate was relatively high, which may have affected the reliability of the Kaplan-Meier estimates. In addition, the clinical information of some patients

was incomplete, which may have influenced the assessment of the independence of the risk score model and reduced the robustness of the survival study. Finally, this study was based on the high-throughput RNA-sequencing profiles and data analyses; therefore, clinical and biological studies are required to validate these findings.

Conclusions

After analyzing TCGA cohorts, we identified a specific signature consisting of 5 lncRNAs for gastric cancer. Despite some limitations, this gastric cancer-specific lncRNA signature is useful for prediction of prognosis.

Acknowledgments

Conflict of interest

The authors declare no competing financial interests.

References:

1. Torre LA, Bray F, Siegel RL et al: Global cancer statistics, 2012. Cancer J Clin, 2015; 65(2): 87–108

2. Maconi G, Manes G, Porro GB: Role of symptoms in diagnosis and outcome of gastric cancer. World J Gastroenterol, 2008; 14(8): 1149–55

3. Saka M, Morita S, Fukagawa T, Katai H: Present and future status of gastric cancer surgery. Jpn J Clin Oncol, 2011; 41(3): 307–13

4. Group G, Oba K, Paoletti X et al: Role of chemotherapy for advanced/recurrent gastric cancer: an individual-patient-data meta-analysis. Eur J Cancer, 2013; 49(7): 1565–77

5. De Angelis R, Sant M, Coleman MP et al: Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE – 5-a population-based study. Lancet Oncol, 2014; 15(1): 23–34

6. Zhao J, Liu Y, Zhang W et al: 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

7. Mercer TR, Dinger ME, Mattick JS: Long non-coding RNAs: insights into functions. Nat Rev Genet, 2009; 10(3): 155–59

8. Geisler S, Coller J: RNA in unexpected places: Long non-coding RNA functions in diverse cellular contexts. Nat Rev Mol Cell Biol, 2013; 14(11): 699–712

9. Kitagawa M, Kitagawa K, Kotake Y et al: Cell cycle regulation by long non-coding RNAs. Cell Mol Life Sci, 2013; 70(24): 4785–94

10. Fatica A, Bozzoni I: Long non-coding RNAs: New players in cell differentiation and development. Nat Rev Genet, 2014; 15(1): 7–21

11. Wapinski O, Chang HY: Long noncoding RNAs and human disease. Trends Cell Biol, 2011; 21(6): 354–61

12. Zhang H, Chen Z, Wang X et al: Long non-coding RNA: a new player in cancer. J Hematol Oncol, 2013; 6: 37

13. Hu Y, Wang J, Qian J et al: Long noncoding RNA GAPLINC regulates CD44-dependent cell invasiveness and associates with poor prognosis of gastric cancer. Cancer Res, 2014; 74(23): 6890–902

14. Song H, Sun W, Ye G et al: Long non-coding RNA expression profile in human gastric cancer and its clinical significances. J Transl Med, 2013; 11: 225

15. Xu MD, Qi P, Weng WW et al: Long non-coding RNA LSINCT5 predicts negative prognosis and exhibits oncogenic activity in gastric cancer. Medicine (Baltimore), 2014; 93(28): e303

16. Zhang E, He X, Yin D et al: Increased expression of long noncoding RNA TUG1 predicts a poor prognosis of gastric cancer and regulates cell proliferation by epigenetically silencing of p57. Cell Death Dis, 2016; 7: e2109

17. Zhou X, Yin C, Dang Y et al: Identification of the long non-coding RNA H19 in plasma as a novel biomarker for diagnosis of gastric cancer. Sci Rep, 2015; 5: 11516

18. Cabanski CR, White NM, Dang HX et al: Pan-cancer transcriptome analysis reveals long noncoding RNAs with conserved function. RNA Biol, 2015; 12(6): 628–42

19. Li J, Han L, Roebuck P et al: TANRIC: An interactive open platform to explore the function of lncRNAs in cancer. Cancer Res, 2015; 75(18): 3728–37

20. Goldman M, Craft B, Swatloski T et al: The UCSC cancer genomics browser: Update 2015. Nucleic Acids Res, 2015; 43(Database issue): D812–17

21. Wettenhall JM, Smyth GK. limmaGUI: A graphical user interface for linear modeling of microarray data. Bioinformatics, 2004; 20(18): 3705–6

22. Lossos IS, Czerwinski DK, Alizadeh AA et al: Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. N Engl J Med, 2004; 350(18): 1828–37

23. Guttman M, Amit I, Garber M et al: Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature, 2009; 458(7235): 223–27

24. Huang da W, Sherman BT, Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc, 2009; 4(1): 44–57

25. Cui M, You L, Ren X et al: Long non-coding RNA PVT1 and cancer. Biochem Biophys Res Commun, 2016; 471(1): 10–14

26. Ashkenazi A: Directing cancer cells to self-destruct with pro-apoptotic receptor agonists. Nat Rev Drug Discov, 2008; 7(12): 1001–12

27. Fuchs Y, Steller H: Programmed cell death in animal development and disease. Cell, 2011; 147(4): 742–58

28. Van Cutsem E, Sagaert X, Topal B et al: Gastric cancer. Lancet, 2016 [Epub ahead of print]

29. Pietrantonio F, De Braud F, Da Prat V et al: A review on biomarkers for prediction of treatment outcome in gastric cancer. Anticancer Res, 2013; 33(4): 1257–66

30. Wang J, Sun J, Wang J et al: Long noncoding RNAs in gastric cancer: functions and clinical applications. Onco Targets Ther, 2016; 9: 681–97

31. Hu Y, Chen HY, Yu CY et al: A long non-coding RNA signature to improve prognosis prediction of colorectal cancer. Oncotarget, 2014; 5(8): 2230–42

32. Zhang XQ, Sun S, Lam KF et al: A long non-coding RNA signature in glioblastoma multiforme predicts survival. Neurobiol Dis, 2013; 58: 123–31

33. Meng J, Li P, Zhang Q et al: A four-long non-coding RNA signature in predicting breast cancer survival. J Exp Clin Cancer Res, 2014; 33: 84

34. Zhou M, Guo M, He D et al: A potential signature of eight long non-coding RNAs predicts survival in patients with non-small cell lung cancer. J Transl Med, 2015; 13: 231

35. Li J, Chen Z, Tian L et al: LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. Gut, 2014; 63(11): 1700–10

36. Sun M, Jin FY, Xia R et al: Decreased expression of long noncoding RNA GAS5 indicates a poor prognosis and promotes cell proliferation in gastric cancer. BMC Cancer, 2014; 14: 319

37. Wang Z, Yan Z, Zhang B et al: Identification of a 5-gene signature for clinical and prognostic prediction in gastric cancer patients upon microarray data. Med Oncol, 2013; 30(3): 678

38. Li X, Zhang Y, Zhang Y et al: Survival prediction of gastric cancer by a seven-microRNA signature. Gut, 2010; 59(5): 579–85

39. Shi J, Qu YP, Hou P: Pathogenetic mechanisms in gastric cancer. World J Gastroenterol, 2014; 20(38): 13804–19

40. Guttman M, Rinn JL: Modular regulatory principles of large non-coding RNAs. Nature, 2012; 482(7385): 339–46

41. Tan P, Yeoh KG: Genetics and molecular pathogenesis of gastric adenocarcinoma. Gastroenterology, 2015; 149(5): 1153–62.e3