Inconsistency in Expression Pattern of a Five-lncRNA Signature As a Potential Diagnostic Biomarker for Gastric Cancer Patients in Bioinformatics and in Vitro

Mahmoud Ghanei

Mashhad University of Medical Sciences

Arash Poursheikhani

Mashhad University of Medical Sciences

Azadeh Aarabi

Mashhad University of Medical Sciences

Negin Taghehchian

Ferdowsi University of Mashhad

Mohammad Reza Abbaszadegan (✉ abbaszadeganmr@mums.ac.ir)

Mashhad University of Medical Sciences

Research Article

Keywords: lncRNA, gastric cancer, TCGA, Signature, diagnostic biomarker

Posted Date: August 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-712731/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

BACKGROUND: Due to the diagnosis of gastric cancer in advance stages as well as its poor prognosis, finding biomarkers is essential.

OBJECTIVE: In this study, using the TCGA RNAseq data of gastric cancer patients, we evaluated the diagnostic value of lncRNAs which had differential expression.

METHODS: we evaluated P value, FDR, log fold change for whole transcripts. Next, by comparison of the RNAseq gene names with total known lncRNA names, we identified differentially expressed lncRNAs. Following, we calculated specificity and sensitivity for lncRNAs from the previous step. For more confirmation, we predicted target genes and performed GO and KEGG signaling pathway analysis. At the end, we examined the reliability and consistency of expression of this signature in three gastric cancer cell lines and one of them in twenty tumoric and tumor adjacent normal tissue samples using qRT-PCR.

RESULTS: Five lncRNAs had proper sensitivity and specificity and had target genes involved in cancer-related signaling pathways; however, they showed different expression patterns in TCGA data and in vitro.

CONCLUSIONS: The results of our study demonstrated that the five-lncRNAs *PART1*, *UCA1*, *DIRC3*, *HOTAIR*, and *HOXA11AS* require more investigation to be confirmed as a diagnostic biomarker in gastric cancer.

1. Background

Gastric cancer (GC) is one of the most important leading causes of cancer-related death worldwide (1). According to the report of Global Cancer Observatory (GLOBOCAN) in 2018, gastric cancer is fifth in incidence and third in mortality among cancers (both sexes and all ages) across the world, but it is the second and the first in incidence and mortality in Iran, respectively (2). Despite advancements in treatment of gastric cancer patients based on surgical approaches and targeted drug chemotherapy, poor prognosis and late manifestation of symptoms act as obstacles on early diagnosis of these patients; therefore, identifying patients at advanced stages of the disease can make most patients untreated (3, 4). Early detection of patients, prediction of outcomes of treatments effectively, identification of new therapeutic targets, a better understanding of tumorigenesis and progression processes are crucial keys for improving the survival rate of GC patients. Therefore, the discovery and development of prognostic and diagnostic biomarkers are essential for the facilitation of early diagnosis, effective prediction of prognosis, resulting in better outcomes in GC patients.

Long noncoding RNAs (lncRNAs) are attributed as a class of noncoding RNAs (ncRNAs) of greater than 200 nucleotides in length, which have a specific expression in various tissues and diseases such as cancers (5–9). lncRNAs are involved in different biological processes including cell development and differentiation, cell cycle arrest, apoptosis, autophagy, cell senescence, chromosome remodeling, X chromosome inactivation, alternative splicing, RNA decay, embryonic stem cells, cancer cell metastasis, drug resistance, etc. (7, 10–14). According to some studies, lncRNAs have a more important function than protein-coding genes in translational and clinical oncology (11, 15). On the other hand, a group of these noncoding transcripts act as a tumor suppressor or oncogene and be dysregulated in various types of cancers (16, 17). Other studies have shown that lncRNAs are correlated with cancer recurrence and poor prognosis although has not been fully elucidated yet in gastric cancer (18, 19). Since lncRNAs participate in various processes in cancers and also easy detection of them, they can be chosen as valuable biomarkers in the diagnosis and prognosis of cancers (20, 21). Several studies have explored the expression of some lncRNAs in GC and introduced several lncRNAs as biomarkers; however, a few numbers have proposed a lncRNA signature, and certainly, there are further lncRNAs that have to be investigated and can be used as practical biomarkers after passing the laboratory steps (22, 23).

In this study, we attempted to perform a comprehensive study using transcriptomic data analysis of patients with gastric cancer that are freely available in The Cancer Genome Atlas database (TCGA) ([HTTP:// cancergenome.nih.gov/](http://cancergenome.nih.gov/)) and to provide a new and effective signature for the diagnosis of gastric cancer patients. With conducting data mining in RNAseq data of GC patients and thereafter determination of differentially expressed genes and calculating AUC of ROC curve, we identified a five-lncRNA signature (*HOTAIR*, *PART1*, *DIRC3*, *HOXA11AS*, *UCA1*) as a novel potential biomarker for diagnosis of GC patients. We also examined the expression of this signature in three gastric cancer cell lines and one of them in twenty tumoric and tumor adjacent normal tissue samples.

2. Materials And Methods

2.1. RNAseq data mining and finding of differentially expressed lncRNAs

The RNAseq data and clinical characteristics were downloaded from TCGA database to the address <https://cancergenome.nih.gov/>. After the data was received, the primary trimming was done and redundant columns and rows were deleted. The RNAseq dataset (STAD) contained 450 transcriptomes, of which 415 were tumoric and 35 were tumor-adjacent normal tissue samples.

The RNAseq data processing was continued using R program. In first, we performed secondary data trimming by virtue determination of the third data quantile and omission of the transcripts that had read counts fewer than 25 in the three quarters of the data. Applying this command resulted in the removal of 4639 transcripts from the original data set with 20531 transcripts. Then, by calling the edgeR library, logFC, logCPM, P-value and FDR were calculated for each transcripts. In the next step, by applying appropriate filtration ($P\text{-value} < 0.05$, $FDR < 0.05$, $1 < \log FC < -1$) on these components, the list of genes with differential expression was obtained. This list included all types of transcripts with differential expression, from which lncRNAs were to be identified. For this purpose, a list of all identified lncRNAs was obtained from the HUGO Gene Nomenclature Committee (HGNC) website and aligned with the list of differential expression genes to identify the lncRNAs with differential expression. The result was 14 lncRNAs, which were nominated for the next step of the analysis.

2.2. ROC curve

This step was done by SPSS software used to determine the sensitivity and specificity of the lncRNAs from previous step, which ultimately led to the selection of 5 lncRNAs as a potential diagnostic biomarker.

2.3. Target genes prediction and functional enrichment analysis

The target genes of the five lncRNA signature were predicted by usage of two online tools include of lncRRsearch and lncRNA2Target. The sum of the identified target genes for these 5 lncRNAs were 238 by the lncRRsearch after the removal of duplicates and 115 by the lncRNA2Target. Furthermore, the functional analysis of the GO annotation, DO and KEGG signaling pathways were performed. P-value less than 0.01 was considered statistically significant.

2.4. Protein–protein interaction network

To predict the relationships between the lncRNAs target genes products in network, the ID number of the lncRNAs target genes and STRING online tool were used. The highest confidence score (a combined score > 0.900) was considered significant. Globe-shaped proteins have been characterized based on their association with other proteins. The target genes with multiple connections to other target genes appear to play important roles in the protein-protein interaction network.

2.5. Statistical analysis

The data were presented as Mean ± SD. ROC curve analysis was conducted by IBM SPSS statistics v25. A p-value < 0.05 was considered statistically significant.

2.6. Samples and cell lines

Normal tissue samples as a control group were obtained from 20 non-cancerous individuals that had been scanned for gastroesophageal diseases via upper endoscopy in Imam Reza hospital, Mashhad, Iran. Healthy control tissue samples were confirmed by pathological examination. The fresh specimens immediately were transferred to RNA later solution and then stored at -70°C before RNA extraction. Written informed consent was obtained from all participants in this study before tissue sampling.

Three gastric cancer cell lines, MKN-45, AGS, and EPG were obtained from the National Cell Bank of Iran, Tehran. For all used cell lines, STR profiling was performed for the determination of cell lines identity; the report was compared with standard cell lines by the usage of matching criteria based on an algorithm that compares the number of shared alleles between two cell lines, expressed as a percentage. Cell lines with ≥ 80% match are considered to be related; derived from common ancestry. Cell lines with between a 55–80% match require further analysis for authentication of relatedness that all three cell lines had an acceptable resemblance.

2.7. RNA extraction and real-time PCR

Total RNA was extracted with RNX-Plus solution (SinaClon, Cat. No. RN7713C, Iran) according to manufacturer's protocol. Next, the purity of the extracted RNAs was evaluated by Spectrophotometer, Biwave II (Biochrom, UK). We also used from Agarose gel electrophoresis to ensure RNA integrity. Next, for eliminating of DNA pollution, we treated the extracted RNAs by DNase1 (Thermo Fisher Scientific, Cat. No. EN0521). Reverse transcription reactions were performed in two steps using the NG dART RT kit (EURx, Cat. No. E0801-03) according to the manufacturer's protocol. QRT-PCR analysis of the expression of the five lncRNAs was performed on a Light Cycler 96 (Roche Life Science, Roche Diagnostics GmbH, Germany) using the suitable primers (Table 1) and RealQ Plus (AMPLIQON, Germany) according to manufacturer's protocol. All reactions were conducted in duplicate form and the expression of the five lncRNAs was normalized to the expression level of the housekeeping gene GAPDH.

Table 1
the primer used for real-time PCR designed by primer3.

Target	Forward Primer	Reverse Primer	Product Size
DIRC3	CTCATCTGTCCGACGAAGCA	CCCTACTGTCCTGGTGGAGA	73
HOTAIR	GGAAGCGAAGGGTTGTGTA	GGCTAGGGCTGGTTTCACTT	175
HOXA11AS	TTTAGAGGCGCTGACATCCG	CTCAGTCGGGTCTTTCCCAG	89
PART1	TCCAGAGCCAGCCAATCACT	TGTCCTTTTCCCCTCCGACA	181
UCA1	GCCAGCCTCAGCTTAATCCA	CCCTGTTGCTAAGCCGATGA	151

3. Results

3.1. Differentially Expressed Genes

The data demonstrated that 3722 genes (1840 up-regulated and 1882 down-regulated) were differentially expressed in STAD. Moreover, 14 lncRNAs (139 up-regulated and 49 down-regulated) were identified that were most deferentially expressed in patients. The data are presented in Tables 2 and 3.

Table 2
Top 20 up and down-regulated Genes.

Top 20 up-regulated genes				
	logFC	logCPM	PValue	FDR
ENPP7	-7.24658	2.654383	4.6E-114	7.3E-110
SLC2A7	-6.4681	-2.81146	1.04E-89	2.37E-86
S100G	-6.41206	0.514534	4.73E-56	2.21E-53
FLG	-6.36187	3.289874	5.1E-111	4.1E-107
GIP	-6.23295	1.270187	1.89E-38	3.8E-36
CLDN22	-6.1317	-3.15007	2.71E-19	5.83E-18
KRT1	-6.10628	1.746563	6.56E-80	7.45E-77
MEP1B	-6.04667	3.490235	2.25E-65	1.79E-62
ZG16	-6.00036	2.506113	1.57E-60	9.99E-58
AQP10	-5.97356	2.003138	7.84E-74	7.33E-71
SLC28A1	-5.88603	0.267127	3.46E-85	6.88E-82
ATP4B	-5.85136	5.777434	5.54E-37	9.89E-35
G6PC	-5.65078	0.41213	5.26E-43	1.39E-40
ATP4A	-5.57913	5.920505	1.5E-27	9.08E-26
CRNN	-5.52021	8.116779	4.7E-13	4.38E-12
CRCT1	-5.49336	4.642107	9.88E-21	2.54E-19
CA7	-5.47359	-0.32813	2.15E-98	6.83E-95
ACER1	-5.3863	-0.21548	2.1E-32	2.34E-30
KPRP	-5.35025	1.760696	1.46E-18	2.87E-17
GYS2	-5.29783	0.208704	1.46E-43	3.94E-41
Top 20 Down-regulated genes				
	logFC	logCPM	PValue	FDR
ENPP7	-7.24658	2.654383	4.6E-114	7.3E-110
SLC2A7	-6.4681	-2.81146	1.04E-89	2.37E-86
S100G	-6.41206	0.514534	4.73E-56	2.21E-53
FLG	-6.36187	3.289874	5.1E-111	4.1E-107
GIP	-6.23295	1.270187	1.89E-38	3.8E-36
CLDN22	-6.1317	-3.15007	2.71E-19	5.83E-18
KRT1	-6.10628	1.746563	6.56E-80	7.45E-77
MEP1B	-6.04667	3.490235	2.25E-65	1.79E-62
ZG16	-6.00036	2.506113	1.57E-60	9.99E-58
AQP10	-5.97356	2.003138	7.84E-74	7.33E-71
SLC28A1	-5.88603	0.267127	3.46E-85	6.88E-82
ATP4B	-5.85136	5.777434	5.54E-37	9.89E-35
G6PC	-5.65078	0.41213	5.26E-43	1.39E-40
ATP4A	-5.57913	5.920505	1.5E-27	9.08E-26
CRNN	-5.52021	8.116779	4.7E-13	4.38E-12
CRCT1	-5.49336	4.642107	9.88E-21	2.54E-19
CA7	-5.47359	-0.32813	2.15E-98	6.83E-95
ACER1	-5.3863	-0.21548	2.1E-32	2.34E-30

Top 20 up-regulated genes				
KPRP	-5.35025	1.760696	1.46E-18	2.87E-17
GYS2	-5.29783	0.208704	1.46E-43	3.94E-41

Table 3
The Differentially expressed lncRNAs.

LncRNAs				
	logFC	logCPM	PValue	FDR
BCAR4	5.410524	-0.55488	5.09E-06	1.76E-05
CDKN2BAS	-2.45409	0.65801	2.32E-19	5.05E-18
DIRC3	-2.05778	-1.47014	2.71E-18	5.12E-17
DLX6AS	3.012512	-0.69193	1.99E-06	7.39E-06
DSCR4	6.162059	-1.17853	2.5E-06	9.13E-06
FAM27B	2.800521	-2.89405	0.029757	0.048477
CDH19	-2.7026	1.429299	6.03E-15	7.21E-14
HOTAIR	5.983826	1.520031	1.99E-22	6.36E-21
HOXA11AS	3.541999	1.706383	1.93E-10	1.26E-09
HULC	6.653741	-0.35465	2.83E-07	1.19E-06
IGF2AS	3.058415	-1.21374	5.47E-06	1.88E-05
PART1	-2.41557	2.244247	9.07E-17	1.38E-15
TTY14	-2.1316	-1.34436	2.58E-07	1.09E-06
UCA1	3.323226	4.911589	3.04E-08	1.47E-07

3.2. ROC curve indicates good performance for the suggested diagnostic signature

For evaluation of the diagnostic power of the fourteen-lncRNA achieved from the RNAseq data analysis by R program, we calculated the AUC of the ROC curve. The higher AUC represents a better performance and an AUC more than 0.7 is acceptable and is considered as good efficiency. P-value less than 0.05 was considered as significant level. According to AUC and p-value, eventually the five lncRNAs (*PART1*, *UCA1*, *DIRC3*, *HOTAIR* and *HOXA11AS*) were selected as the candidate diagnostic signature. These selected lncRNAs had good potential sensitivity and specificity for the diagnosis of gastric cancer patients (Fig. 1). The expression pattern of these lncRNAs have been shown in Fig. 2.

3.3. The target gene prediction

The target genes of the diagnostic lncRNAs signature were predicted by LncRRlsearch (24) and LncRNA2Target (25) online tools. The number of the target genes identified for lncRNAs *PART1*, *UCA1*, *DIRC3*, *HOTAIR* and *HOXA11AS* by the LncRRlsearch database were 100, 100, 33, 100 and 100 respectively (this database represents the top 100 genes identified for the lowest sum of energy) and by the LncRNA2Target database were 0, 55, 0, 68 and 10, respectively. In total, 433 target genes from the LncRRlsearch database and 133 target genes from the LncRNA2Target database were identified for this signature. After deleting duplicates and merging two lists, a list with 353 target genes was generated.

3.4. GO annotation, Disease Ontology and KEGG signaling pathways

For more investigation about the potential biological function and the mechanism of the suggested diagnostic signature, we took advantage of GO annotation, Diseases ontology and KEGG signaling pathway analysis for 353 target genes from the previous step. The GO annotation, Diseases ontology, and KEGG signaling pathway analysis were conducted and visualized by R software (ggplot2 package). P-value less than 0.01 was considered as cut off point. From the examination of these results, we found many functions that are related to EMT process, RNA silencing mechanisms, chromatin remodeling, molecular binding, gene transcription and cell proliferation. Therefore, it can be deduced that this suggested diagnostic signature may be related to gene expression and molecular and cellular functions. The results of KEGG signaling pathway analysis and disease ontology (DO) indicate activation of cancer-related pathways such as gastric cancer, prostate cancer, small cell lung cancer, pancreatic cancer, breast cancer, bladder cancer, colorectal cancer, melanoma, glioma, connective tissue cancer, bone cancer, renal carcinoma, etc. (Fig. 3).

3.5. Functional predictions and PPI network construction

To predict and visualize the protein-protein interactions among the target genes of the five-lncRNA signature, we used from STRING (27) software. For this purpose, to obtain protein-protein interaction (PPI) data, the target genes were first introduced in the STRING website. Next, the cut off greater than 0.900 for a combined score of PPIs was considered as appropriate criteria for selection and to construct PPI networks. In this network, unconnected nodes are not displayed. Of the 353 target genes identified, 252 genes were involved in constructing nodes of this network (Fig. 4). The genes with connections/interactions

more than 7 (degree > 7) were filtered. STRING data demonstrated that these lncRNAs contributed to several canonical signaling pathways related proteins such as cyclins and related *CDKs*, *TWIST1*, *CDH1*, *MMPs*, *TP53*, *RB1*, *WNT7A*, *GSK3B*, *ATM*, *PTEN*, *BCR*, *ERBB2*, *VEGFR*, *SMAD*, *TGFB2*, *MAPK1*, *FGF1*, *AKT1* and *HIF1A*.

3.6. Functional analysis via co-expressed target genes

At the beginning of this study and before selecting the appropriate lncRNAs as biomarkers, a list of genes with differential expression of TCGA data was obtained and then a list of target genes was predicted by online tools for this signature. In order to probe the functional value of the signature in the gastric cancer (stomach adenocarcinoma), we searched the predicted target genes that were co-expressed with our suggested signature and we found a list of 212 genes with this characteristic. In this new list, there are many master genes such as *CDKs*, *TWIST1*, *CDH1*, *MMPs*, *TP53*, *RB1*, *WNT7A*, *GSK3B*, *ATM*, *PTEN*, *BCR*, *ERBB2*, *VEGFR*, *SMAD*, *TGFB2*, *MAPK1*, *FGF1*, *AKT1* and *HIF1A* that involved in cancer-related pathways and therefore, dysregulated in many cancers.

3.7. Real-time PCR verification

We investigated the expression level of the suggested five-lncRNA signature in the three gastric cancer cell lines and 20 healthy tissue samples. We couldn't confirm the similar expression pattern for these lncRNAs in the cell lines and TCGA RNAseq data of gastric cancer patients using real-time PCR. We used from paired t-test as a proper statistical analysis for the compare expression level of the five-lncRNA signature in three gastric cancer cell lines with 20 healthy tissue samples. We found that *PART1* and *DIRC3* had no detectable expression in none of the three cell lines and normal tissue samples. In the other hand, *UCA1* was expressed in whole samples. Two lncRNAs, *HOTAIR* and *HOXA11AS*, were similarly expressed only in the EPG cell line and normal tissue samples.

We also examined the expression level of *UCA1* in twenty tumoric and tumor-adjacent normal tissue samples with confirmed expression level greater than 2 (logFC > 2) for *TWIST1*. This study showed that there is no significant correlation between expression level of *UCA1* and *TWIST1* in tissue samples which is in contrast to our expectations for *UCA1* expression.

4. Discussion And Conclusions

In our investigation, we used the TCGA RNAseq data to introduce a lncRNA signature which can be utilized as diagnostic biomarker. Our bioinformatics data was illustrated that lncRNAs *PART1*, *UCA1*, *DIRC3*, *HOTAIR* and *HOXA11AS* have more differential expression in the tumoric tissues vs to the normal counterpart margins. Moreover, ROC curve analysis showed that these lncRNAs have significant sensitivity and specificity (diagnostic) values.

Furthermore, we investigated the role of the lncRNAs through the enrichments and real-time PCR.

It has been demonstrated that down-regulation of lncRNA *PART1* blocked cell proliferation and promotes apoptosis in bladder cancer (28). Furthermore, exosome-mediated transfer of lncRNA *PART1* can induced chemotherapy resistance in ESCC by competing endogenous miRNA (29). *PART1* promotes tumorigenesis by *miR-143* in colorectal cancer (30). *Leung GK et al.* declared that a long non-coding RNA signature can predict prognosis of the glioblastoma multiform patients (31).

Overexpression of lncRNA *UCA1* has been suggested that increases cell growth and chemo-resistance by inhibiting *miR-513a-5p* in retinoblastoma cells (32). In addition, *UCA1* overexpression is associated with poor survival rate in patients with digestive system malignancies (33). *Ma C et al.* demonstrated that high expression of serum *UCA1* can be consider as a potential biomarker for clinical diagnosis of gastric cancer (34). It has been suggested that *UCA1* confers endocrino-therapy resistance through EZH2/p21 axis and the PI3K/AKT signaling pathway in breast cancer (35). *UCA1* can promote carcinogenesis by Wnt signaling pathway in papillary thyroid carcinoma (36). Downregulation of *UCA1* has been showed that enhances the radio-sensitivity and inhibits cell migration by suppression of EMT in colorectal cancer cells (37). Moreover, *Yuan D et al.* found that the Knockdown of *UCA1* increases cisplatin sensitivity in tongue squamous cell carcinoma cells (38).

Previous investigations have reported that lncRNA *DIRC3* expression had impact on carcinogenesis. This lncRNA has been regulated by *MITF-SOX10* in melanoma tumor (39). *Shao Y et al.* showed that *DIRC3* and near *NABP1* genetic polymorphisms are associated with poor prognosis in laryngeal squamous cell carcinoma patient (40).

lncRNA *HOTAIR* has been demonstrated that can be serve as a prognostic and diagnostic biomarker in cancers. It promotes tumor progression via sponging *miR-217-GPC5* axis in gastric cancer (41). *HOTAIR* mediates the switching of histone H3 lysine 27 acetylation to methylation, to promote EMT process in gastric cancer (42). Moreover, it has been reported that *HOTAIR* rs17720428 SNP is correlated with the risk and prognosis of gastric cancer in the Chinese Han population. *HOTAIR* modulates *KLF12* to regulate gastric cancer progression via PI3K/ATK signaling pathway by sponging *miR-618* (44). It has been reported that lncRNAs such as *H19*, *HOTAIR*, *UCA1* and *PVT1* could be serve as potential diagnostic and prognostic biomarkers in patients with gastric cancer (45). *Chen LC et al.* demonstrated that high *HOTAIR* expression promote proliferation and metastasis in gastric cancer via *miR-126/CXCR4* axis (46).

Chen P et al. exhibited that lncRNA *HOXA11-AS* acted as a ceRNA to promote cisplatin resistance of human LUAD cells via the *miR-454-3p/Stat3* axis (47). *Li W et al.* showed which increased expression of *HOXA11-AS* is a risk factor for poor clinical outcomes in numerous tumors and may act as a novel biomarker for poor prognosis and metastasis in cancers (49). *Hu XF et al.* demonstrated the oncogenic role of *HOXA11AS* in breast cancer, providing novel clues for the future clinical diagnosis and treatment of early stage breast cancer patients (50). *Hu CP et al.* showed that up-regulation of lncRNA *HOXA11-AS* predicted a poor prognosis and lncRNA *HOXA11-AS* promoted cell epithelial-mesenchymal transition (EMT) by inhibiting *miR-200b* expression in NSCLC (51). *Sun M et al.* declared that *HOXA11-AS* not only could promote GC cells migration and invasion in vitro, but also promotes GC cells metastasis in vivo, at least in part, by regulating β -catenin and *KLF2* (52). lncRNA *HOXA11-AS* was showed that able to distinguish CRC tissue from non-cancerous tissue, and CRC tissue with lymph node metastasis from CRC without lymph node metastasis (53).

To clarify the molecular function of this suggested five-lncRNA signature, we predicted the target genes and the corresponding pathways using GO annotation, disease ontology (DO) and KEGG signaling pathway analysis. The results showed the participation of this signature in essential biological processes such as cell division, transcription regulation, change in expression of growth factors and enriched KEGG pathways including PI3k-Akt signaling pathways, p53 signaling pathways and pluripotent stem cell signaling pathways. In order to further investigation of the proposed lncRNA signature, the protein-protein interactions of the target genes were predicted.

The current study indicated inconsistency in the results of real-time PCR in the cell lines and TCGA RNAseq data. LncRNAs *PART1* and *DIRC3* had no detectable expression in any cell lines. These two lncRNAs were down-regulated in tumoric TCGA RNAseq samples vs normal margin; lack of expression of them in the cell lines is probably due to very low or no expression of them which emanated from cumulative mutations that cell lines bear in passing time, presence of different subclones in culture and so, their complex and unclear interactions, different circumstance of cells in culture vs body, and other causes that are still unclear. The function of lncRNAs *PART1* and *DIRC3* is unclear in gastric cancer; therefore, future studies should focus on these lncRNAs and clarify their function in gastric cancer. In the other hand, lncRNA *UCA1* was expressed in all three cell lines; however, its expression was differed from TCGA data. Of the probable reasons for its expression in all three cell lines, could be its more important role in gastric carcinogenesis, more stability and more expression; however, deviation in the expression pattern with TCGA data is among the cases that are very difficult to comment on. Nevertheless, *HOTAIR* and *HOXA11AS* were expressed just in EPG cell line and had no detectable expression in other two cell lines. This can be explained with role of these lncRNAs in EPG cell line but in no others, because of different genetic context of three cell lines. Overall, the exact interpretation of these discrepancies between the results of bioinformatics and laboratory studies requires further study, and here we have only stated a series of hypotheses.

The activation of PI3K-Akt pathway is reported in many malignancies. The inhibition of the PI3K-Akt pathway can induce apoptosis and decrease cell division via negative regulation of *Plk1* both in vitro and in vivo (56, 57). One of the most important gene involved in this pathway is *TWIST1*. This gene normally is overexpressed in many cancers and promote metastasis using activation of EMT process. *TWIST1* was predicted as an important target gene for *UCA1* and also was existed in differential expression genes list resulted from TCGA data analysis. In order to examination of correlation between *TWIST1* and *UCA1*, the lncRNA that had been expressed in all three cell lines, we performed real-time PCR for *UCA1* in twenty tumoric and tumor-adjacent normal tissue samples with confirmed over-expression of *TWIST1* (logFC > 2). This examination showed that *UCA1* was upregulated and downregulated in 8 and 3 samples, respectively. In other 9 tissue samples, *UCA1* did not show altered expression. Altogether, evaluation of *UCA1* using paired t-test did not show a significant statistical relationship in twenty tissue samples; This was another important contradiction that we encountered during this study. Discrimination in sampling (selection of samples with high *TWIST1* expression) seems to be one of the reasons for this unexpected result; However, based on bioinformatics analysis, our sampling was correct and the results should have been different.

In total, like any other studies, our study had its limitations. First, the data used for this study to introduce a diagnostic signature were extracted from a single database (TCGA); the use of more databases such as GEO certainly makes the results more reliable. Second, due to the use of cell lines in this study for investigation of suggested signature expression and the differences between these cells in terms of growth conditions and cumulative mutations, as well as the low number of cell lines used, it is necessary to examine the expression of these 5 lncRNAs in and more tissue samples.

In conclusively, we could not present a five-lncRNA signature with diagnostic potential for gastric cancer because of some identified contradictions in bioinformatics and laboratory study. More investigations should be performed for ultimate validation or rejection of this signature; investigations that cover our faults using large sample size, examination of these result in tissue samples and also further exploration in biological and molecular mechanism of suggested five-lncRNA signature in gastric cancer progression.

Declarations

- Ethics approval and consent to participate: The ethics committee of Mashhad University of Medical Sciences approved the study and the committee's reference number is 1398.522
- Consent for publication: Not applicable
- Availability of data and materials: The datasets generated and analysed during the current study are available in the FireBrowse repository via link: http://gdac.broadinstitute.org/runs/stddata__2016_01_28/data/STAD/20160128/gdac.broadinstitute.org_STAD.Merge_maseq_illumina_hiseq_maseq
- Competing interests: The authors declare that they have no competing interests
- Funding: This study was funded by The Vice Chancellor for Research at Mashhad University of Medical Sciences by no.
- Authors' contributions:

CONCEPTION: Ghanei M., Poursheikhani A., Abbaszadegan MR.

INTERPRETATION OR ANALYSIS OF DATA: Ghanei M., Poursheikhani A., Abbaszadegan MR., Aarabi A., Taghehchian N.

PREPARATION OF THE MANUSCRIPT: Ghanei M., Poursheikhani A.

REVISION FOR IMPORTANT INTELLECTUAL CONTENT: Abbaszadegan MR.

SUPERVISION: Abbaszadegan MR.

- Acknowledgements: "The results shown here are in whole or part based upon data generated by the TCGA Research Network: <https://www.cancer.gov/tcga>."

References

1. Sugano K. Screening of gastric cancer in Asia. *Best Pract Res Clin Gastroenterol*. 2015;29(6):895–905.
2. GLOBCAN 2018. CANCER TODAY [Internet]. GLOBCAN. 2018. Available from: https://gco.iarc.fr/today/online-analysis-pie?v=2018&mode=cancer&mode_population=continents&population=900&populations=900&key=total&sex=0&cancer=39&type=0&statistic=5&prevalence=C
3. Maconi G, Manes G, Porro GB. Role of symptoms in diagnosis and outcome of gastric cancer. *World J Gastroenterol WJG*. 2008;14(8):1149.
4. 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.
5. Yin Y, Yan P, Lu J, Song G, Zhu Y, Li Z, et al. Opposing roles for the lncRNA haunt and its genomic locus in regulating HOXA gene activation during embryonic stem cell differentiation. *Elsevier [Internet]*. [cited 2020 Apr 7]; Available from: <https://www.sciencedirect.com/science/article/pii/S1934590915001186>
6. Kitagawa M, Kitagawa K, Kotake Y, Niida H, Ohhata T. Cell cycle regulation by long non-coding RNAs. Vol. 70, *Cellular and Molecular Life Sciences*. 2013. p. 4785–94.
7. Fatica A, Genetics IB-NR, 2014 undefined. Long non-coding RNAs: new players in cell differentiation and development. *nature.com [Internet]*. [cited 2020 Apr 7]; Available from: <https://www.nature.com/articles/nrg3606>
8. Zhang H, Chen Z, Wang X, Huang Z, He Z, Chen Y. Long non-coding RNA: A new player in cancer. Vol. 6, *Journal of Hematology and Oncology*. 2013.
9. Young RS, Ponting CP. Identification and function of long non-coding RNAs. *Essays Biochem*. 2013;54(1):113–26.
10. Quan M, Chen J, Zhang D. Exploring the Secrets of Long Noncoding RNAs. *OPEN ACCESS Int J Mol Sci [Internet]*. 2015 [cited 2020 Apr 7];16:16. Available from: www.mdpi.com/journal/ijms
11. Evans J, ... FF-TJ of clinical, 2016 undefined. The bright side of dark matter: lncRNAs in cancer. *Am Soc Clin Investig [Internet]*. [cited 2020 Apr 7]; Available from: <https://www.jci.org/articles/view/84421>
12. Spizzo R, Almeida MI, Colombatti A, Calin GA. Long non-coding RNAs and cancer: a new frontier of translational research? *Oncogene*. 2012;31(43):4577–87.
13. Kino T, Hurt DE, Ichijo T, Nader N, Chrousos GP. Noncoding RNA gas5 is a growth arrest–and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal*. 2010;3(107):ra8–ra8.
14. Stower H. Epigenetics: X inactivation by titration. *Nat Rev Genet*. 2013;14(8):518.
15. Ren W, Zhang J, Li W, Li Z, Hu S, Suo J, et al. A tumor-specific prognostic long non-coding RNA signature in gastric cancer. *Med Sci Monit Int Med J Exp Clin Res*. 2016;22:3647.
16. Zhou Y, Zhong Y, Wang Y, Zhang X, Batista DL, Gejman R, et al. Activation of p53 by MEG3 non-coding RNA. *J Biol Chem*. 2007;282(34):24731–42.
17. Rinn JL. lncRNAs: linking RNA to chromatin. *Cold Spring Harb Perspect Biol*. 2014;6(8):a018614.
18. Kim HS, Minna JD, White MA. GWAS meets TCGA to illuminate mechanisms of cancer predisposition. *Cell*. 2013;152(3):387–9.
19. Zhang Y, Li H, Zhang W, Che Y, Bai W, Huang G. LASSO-based Cox-PH model identifies an 11-lncRNA signature for prognosis prediction in gastric cancer. *Mol Med Rep*. 2018;18(6):5579–93.
20. Zhou X, Yin C, Dang Y, Ye F, Zhang G. Identification of the long non-coding RNA H19 in plasma as a novel biomarker for diagnosis of gastric cancer. *Sci Rep*. 2015;5:11516.
21. Cabanski CR, White NM, Dang HX, Silva-Fisher JM, Rauck CE, Cicka D, et al. Pan-cancer transcriptome analysis reveals long noncoding RNAs with conserved function. *RNA Biol*. 2015;12(6):628–42.
22. Fang X, Pan H, Leng R, Ye D. Long noncoding RNAs: novel insights into gastric cancer. *Cancer Lett*. 2015;356(2):357–66.
23. Miao Y, Sui J, Xu S-Y, Liang G-Y, Pu Y-P, Yin L-H. Comprehensive analysis of a novel four-lncRNA signature as a prognostic biomarker for human gastric cancer. *Oncotarget*. 2017;8(43):75007.
24. Tsukasa Fukunaga, Junichi Iwakiri YO and MH. LncRRlsearch. 2019; Available from: <http://rtools.cbrc.jp/LncRRlsearch/index.cgi>
25. Liang Cheng, Pingping Wang, Rui Tian, Song Wang, Qinghua Guo, Meng Luo, Wenyang Zhou, Guiyou Liu, Huijie Jiang and QJ. LncRNA-Target [Internet]. Lncrna2target V2.0: A Comprehensive Database for Target Genes of Lncrnas in Human and Mouse. 2018. Available from: <http://123.59.132.21/lncrna2target/search.jsp>
26. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B IT. Cytoscape: a software environment for integrated models of biomolecular interaction networks [Internet]. *Genome Research*. 2003. p. 13(11):2498-504. Available from: <https://cytoscape.org/index.html>
27. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ von MC. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets [Internet]. *Nucleic acids research*. p. 47:D607-613. Available from: <https://string-db.org/>
28. Hu X, Feng H, Huang H, Gu W, Fang Q, Xie Y, et al. Downregulated Long Noncoding RNA PART1 Inhibits Proliferation and Promotes Apoptosis in Bladder Cancer. *Technol Cancer Res Treat*. 2019 Jan;18:1533033819846638.
29. Kang M, Ren M, Li Y, Fu Y, Deng M, Li C. Exosome-mediated transfer of lncRNA PART1 induces gefitinib resistance in esophageal squamous cell carcinoma via functioning as a competing endogenous RNA. *J Exp Clin Cancer Res*. 2018 Jul;37(1):171.
30. Hu Y, Ma Z, He Y, Liu W, Su Y, Tang Z. PART-1 functions as a competitive endogenous RNA for promoting tumor progression by sponging miR-143 in colorectal cancer. *Biochem Biophys Res Commun*. 2017 Aug;490(2):317–23.

31. Zhang X-Q, Sun S, Lam K-F, Kiang KM-Y, Pu JK-S, Ho AS-W, et al. A long non-coding RNA signature in glioblastoma multiforme predicts survival. *Neurobiol Dis.* 2013 Oct;58:123–31.
32. Yang L, Zhang L, Lu L, Wang Y. lncRNA UCA1 Increases Proliferation and Multidrug Resistance of Retinoblastoma Cells Through Downregulating miR-513a-5p. *DNA Cell Biol.* 2020 Jan;39(1):69–77.
33. Shi F-T, Chen L-D, Zhang L-F. Long Noncoding RNA UCA1 Overexpression Is Associated with Poor Prognosis in Digestive System Malignancies: A Meta-analysis. *Curr Med Sci.* 2019 Oct;39(5):694–701.
34. Shan L, Liu C, Ma C. High Expression of Serum UCA1 may be a Potential Biomarker for Clinical Diagnosis of Gastric Cancer. *Clin Lab.* 2019 Sep;65(9).
35. Li Z, Yu D, Li H, Lv Y, Li S. Long noncoding RNA UCA1 confers tamoxifen resistance in breast cancer endocrinotherapy through regulation of the EZH2/p21 axis and the PI3K/AKT signaling pathway. *Int J Oncol.* 2019 Mar;54(3):1033–42.
36. Lu H-W, Liu X-D. UCA1 promotes papillary thyroid carcinoma development by stimulating cell proliferation via Wnt pathway. *Eur Rev Med Pharmacol Sci.* 2018 Sep;22(17):5576–82.
37. Yang X, Liu W, Xu X, Zhu J, Wu Y, Zhao K, et al. Downregulation of long noncoding RNA UCA1 enhances the radiosensitivity and inhibits migration via suppression of epithelial-mesenchymal transition in colorectal cancer cells. *Oncol Rep.* 2018 Sep;40(3):1554–64.
38. Wang J, Li L, Wu K, Ge W, Zhang Z, Gong L, et al. Knockdown of long noncoding RNA urothelial cancer-associated 1 enhances cisplatin chemosensitivity in tongue squamous cell carcinoma cells. *Pharmazie.* 2016 Oct;71(10):598–602.
39. Coe EA, Tan JY, Shapiro M, Louphrasitthiphol P, Bassett AR, Marques AC, et al. The MITF-SOX10 regulated long non-coding RNA DIRC3 is a melanoma tumour suppressor. *PLoS Genet.* 2019 Dec;15(12):e1008501.
40. Shen Z, Ren W, Bai Y, Chen Z, Li J, Li B, et al. DIRC3 and near NABP1 genetic polymorphisms are associated laryngeal squamous cell carcinoma patient survival. *Oncotarget.* 2016 Nov;7(48):79596–604.
41. Dong X, He X, Guan A, Huang W, Jia H, Huang Y, et al. Long non-coding RNA Hotair promotes gastric cancer progression via miR-217-GPC5 axis. *Life Sci.* 2019 Jan;217:271–82.
42. Song Y, Wang R, Li L-W, Liu X, Wang Y-F, Wang Q-X, et al. 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.* 2019 Jan;54(1):77–86.
43. Li Y, Zhang Z, Yang Y, Ma J. Long Noncoding RNA HOX Transcript Antisense RNA Gene rs17720428 Single Nucleotide Polymorphism Is Associated with Gastric Cancer Risk and Prognosis. *Genet Test Mol Biomarkers.* 2020 Jan;24(1):38–46.
44. Xun J, Wang C, Yao J, Gao B, Zhang L. Long Non-Coding RNA HOTAIR Modulates KLF12 to Regulate Gastric Cancer Progression via PI3K/ATK Signaling Pathway by Sponging miR-618. *Onco Targets Ther.* 2019;12:10323–34.
45. Fattahi S, Kosari-Monfared M, Golpour M, Emami Z, Ghasemian M, Nouri M, et al. lncRNAs as potential diagnostic and prognostic biomarkers in gastric cancer: A novel approach to personalized medicine. *J Cell Physiol.* 2020 Apr;235(4):3189–206.
46. Xiao J, Lai H, Wei S-H, Ye Z-S, Gong F-S, Chen L-C. lncRNA HOTAIR promotes gastric cancer proliferation and metastasis via targeting miR-126 to active CXCR4 and RhoA signaling pathway. *Cancer Med.* 2019 Nov;8(15):6768–79.
47. Zhao X, Li X, Zhou L, Ni J, Yan W, Ma R, et al. lncRNA HOXA11-AS drives cisplatin resistance of human LUAD cells via modulating miR-454-3p/Stat3. *Cancer Sci.* 2018 Oct;109(10):3068–79.
48. Chen D, Sun Q, Zhang L, Zhou X, Cheng X, Zhou D, et al. The lncRNA HOXA11-AS functions as a competing endogenous RNA to regulate PADI2 expression by sponging miR-125a-5p in liver metastasis of colorectal cancer. *Oncotarget.* 2017 Sep;8(41):70642–52.
49. Li N, Yang M, Shi K, Li W. Long non-coding RNA HOXA11-AS in human cancer: A meta-analysis. *Clin Chim Acta.* 2017 Nov;474:165–70.
50. Su J-C, Hu X-F. Long noncoding RNA HOXA11AS promotes cell proliferation and metastasis in human breast cancer. *Mol Med Rep.* 2017 Oct;16(4):4887–94.
51. Chen J-H, Zhou L-Y, Xu S, Zheng Y-L, Wan Y-F, Hu C-P. Overexpression of lncRNA HOXA11-AS promotes cell epithelial-mesenchymal transition by repressing miR-200b in non-small cell lung cancer. *Cancer Cell Int.* 2017;17:64.
52. Liu Z, Chen Z, Fan R, Jiang B, Chen X, Chen Q, et al. Over-expressed long noncoding RNA HOXA11-AS promotes cell cycle progression and metastasis in gastric cancer. *Mol Cancer.* 2017 Apr;16(1):82.
53. Li T, Xu C, Cai B, Zhang M, Gao F, Gan J. Expression and clinicopathological significance of the lncRNA HOXA11-AS in colorectal cancer. *Oncol Lett.* 2016 Nov;12(5):4155–60.
54. Liu Y, Zhang Y-M, Ma F-B, Pan S-R, Liu B-Z. Long noncoding RNA HOXA11-AS promotes gastric cancer cell proliferation and invasion via SRSF1 and functions as a biomarker in gastric cancer. *World J Gastroenterol.* 2019 Jun;25(22):2763–75.
55. Sun M, Nie F, Wang Y, Zhang Z, Hou J, He D, et al. lncRNA HOXA11-AS Promotes Proliferation and Invasion of Gastric Cancer by Scaffolding the Chromatin Modification Factors PRC2, LSD1, and DNMT1. *Cancer Res.* 2016 Nov;76(21):6299–310.
56. Spankuch B, Heim S, Kurunci-Csacsko E, Lindenau C, Yuan J, Kaufmann M, et al. Down-regulation of Polo-like kinase 1 elevates drug sensitivity of breast cancer cells in vitro and in vivo. *Cancer Res.* 2006 Jun;66(11):5836–46.
57. Choi BH, Pagano M, Dai W. Plk1 protein phosphorylates phosphatase and tensin homolog (PTEN) and regulates its mitotic activity during the cell cycle. *J Biol Chem.* 2014 May;289(20):14066–74.
58. Hadley Wickham. ggplot2: Elegant Graphics for Data Analysis [Internet]. Springer-Verlag New York; 2016. p. 978-3-319-24277–4. Available from: <https://ggplot2.tidyverse.org>

Figures

Figure 1
ROC curve analysis of the lncRNAs. A. HOTAIR, B. HOXA11AS, C. UCA1, D. DIRC3, E. PART1.

Figure 2
The expression pattern of suggested lncRNA signature.

Figure 3

GO, DO and KEGG pathway enrichment analysis of the differentially expressed genes (Top 20 GO enrichment are presented). A. GO, B. DO, C. KEGG pathway

Figure 4

Protein-protein interaction (PPI) network of the differentially expressed genes of STAD (score > 0.9000).