Article

Tissue-Specific Down-Regulation of the Long Non-Coding RNAs PCAT18 and LINC01133 in Gastric Cancer Development

Kobra Foroughi [1], Mohammad Amini [1] , Amir Atashi [2], Habibollah Mahmoodzadeh [3], Ute Hamann [4] and Mehdi Manoochehri [1,4,*]

[1] School of Medicine, Shahroud University of Medical Sciences, Shahroud 3614773955, Iran;
   k.foroughi@shmu.ac.ir (K.F.); mohammadamini0624@gmail.com (M.A.)
[2] Department of Medical Laboratory Sciences, School of Allied Medical Sciences, Shahroud University
   of Medical Sciences, Shahroud 3614773955, Iran; atashia@shmu.ac.ir
[3] Department of Surgery, Cancer Institute, Tehran University of Medical Sciences, Tehran 6718743161, Iran;
   hmahmoodzadeh@tums.ac.ir
[4] Molecular Genetics of Breast Cancer (B072), German Cancer Research
   Center (DKFZ), 69120 Heidelberg, Germany; u.hamann@dkfz-heidelberg.de
[*] Correspondence: m.manoochehri@dkfz.de; Tel.: +49-6221-422347; Fax: +49-6221-424722

check for
updates

Abstract: Gastric cancer (GC) is the fifth most common cancer and the third most frequent cause of cancer deaths worldwide. The high death rate associated with GC, and lack of appropriate biomarkers for diagnosis, prognosis, and treatment emphasize the need for identification of novel molecules. Given the emerging roles for long non-coding RNAs (lncRNAs) in cancer development, we studied novel lncRNA candidates involved in gastric carcinogenesis. LncRNA candidate discovery was performed using analyses of available datasets and literature. Validation was done using an internal sample set of GC/normal tissues, and external independent datasets. Network analysis and functional annotation of co-expressed protein coding genes were performed using the weighted gene correlation network analysis (WGCNA) and ingenuity pathway analysis. Two novel lncRNAs, *PCAT18* and *LINC01133*, associated with GC development were identified by analysis of the discovery Gene Expression Omnibus (GEO) datasets. The down-regulation of these genes in GC tissues was successfully validated internally and externally. The results showed a tissue-specific down-regulation of *PCAT18* and *LINC01133* in gastrointestinal tissues. WGCNA and ingenuity pathway analyses revealed that the genes co-expressed with the two lncRNAs were mostly involved in metabolic pathways and networks of gastrointestinal disease and function. Our findings of a tissue-specific down-regulation of *PCAT18* and *LINC01133* in gastric and other gastrointestinal cancers imply that these lncRNAs may have a tumor suppressive function in the development of these tumor entities. The two lncRNA biomarkers may contribute to a better understanding of the complex mechanisms of gastric carcinogenesis.

Keywords: gastric cancer; long noncoding RNA; gene expression; tumor suppressor; GEO and TCGA databases

1. Introduction

Gastric cancer (GC)/stomach cancer is the fifth most common cancer in the world with almost one million new cases reported in 2012 [1]. More than 70% of GCs occur in developing countries and half the world's total occurs in Eastern Asia. GC is the third leading cause of cancer-related death with the highest estimated mortality rate observed in Eastern Asia. Despite advances in diagnosis,

approximately half of all GC patients are diagnosed at an advanced stage and have a poor five -year survival rate of less than 20% [2,3]. Furthermore, surgery and cytotoxic chemotherapy have limited value in advanced disease and markers for targeted therapy are scarce. The etiology of GC is multifactorial involving genetic, epigenetic, and environmental risk factors including *Helicobacter pylori* infection, dietary, lifestyle, and other factors [4–6].

Long non-coding RNAs (lncRNAs) participate in various biological processes including cell proliferation, cell differentiation, migration, invasion, and apoptosis, mainly by regulating of gene expression at the epigenetic, transcriptional, post-transcriptional, and translational levels [7,8]. They execute their various biological functions through different mechanisms as transcriptional enhancers, signals, decoys, guides, or as scaffolds for their interacting proteins; DNA or RNA molecules. LncRNAs are also crucial players in the development and progression of cancer [9–11]. A dysregulation of these genes was reported in a broad spectrum of tumors [12–14], which correlated with the stage and prognosis of several tumor types, and was linked to chemotherapy resistance and targeted therapy [15].

Accumulating evidence has demonstrated that many lncRNAs are also dysregulated in GC and closely related to tumorigenesis, metastasis, prognosis, and diagnosis. Several oncogenic and tumor-suppressive lncRNA genes involved in the development of the disease have been identified [16]. Expression levels of the widely studied oncogenic lncRNAs *H19* and *HOTAIR* have been shown to be associated with different cellular phenotypes and clinical characteristics [17,18]. While expression of the oncoge *BANCR* correlated with tumor invasion depth, metastasis (M) status, lymph node status, and clinical stage in GC patients [19]. Of the tumor-suppressive lncRNAs, *MEG3* correlated with stage, tumor invasion depth, and tumor size [20], whereas expression of *GAS5* correlated with prognosis [21].

The high death rate and the lack of appropriate biomarkers for diagnosis, prognosis, and treatment of GC demand the identification of novel molecules. Given the promising role of lncRNAs in cancer development, the present study aimed to identify novel lncRNAs associated with gastric carcinogenesis. For this purpose, we used lncRNA expression data of GC and normal tissues acquired from public repositories, in addition to a sample set of 25 GC and paired normal tissues. To predict the function of differentially expressed lncRNA candidates, weighted gene correlation network analysis (WGCNA) co-expression network analysis was performed followed by functional pathway enrichment analysis of the co-expressed genes.

2. Results

2.1. Identification of LncRNA Candidates Associated with Gastric Cancer Using Gene Expression Omnibus (GEO) Datasets and Literature Data

An overview of the analysis pipeline is presented in Figure 1A. Five lncRNA candidates were selected; two, *EWSAT1* and *GAS6-AS1*, were selected based on a comprehensive literature search on novel lncRNAs with no previous association with GC development, but deregulated in other tumor entities (Table S1); and three, *PCAT18*, *DANCR*, and *LINC01133*, were selected based on the analysis of GC datasets from GEO.

LncRNA expression data analysis of the four GEO discovery datasets resulted in the identification of 13 differentially expressed lncRNAs (four-GEO candidate set; Table 1 and Figure 1A). Of these, 12 lncRNAs were down-regulated (all $p < 0.05$) in GC tissues and one was upregulated ($p < 0.05$) compared to normal tissues. The number of common and unique lncRNAs across the different datasets is shown in the Venn diagram (Figure 1B). Expression of three lncRNAs, *PCAT18*, *DANCR*, and *LINC01133*, was validated in three other GEO datasets (validation 1; Figure 2 and Table S2).
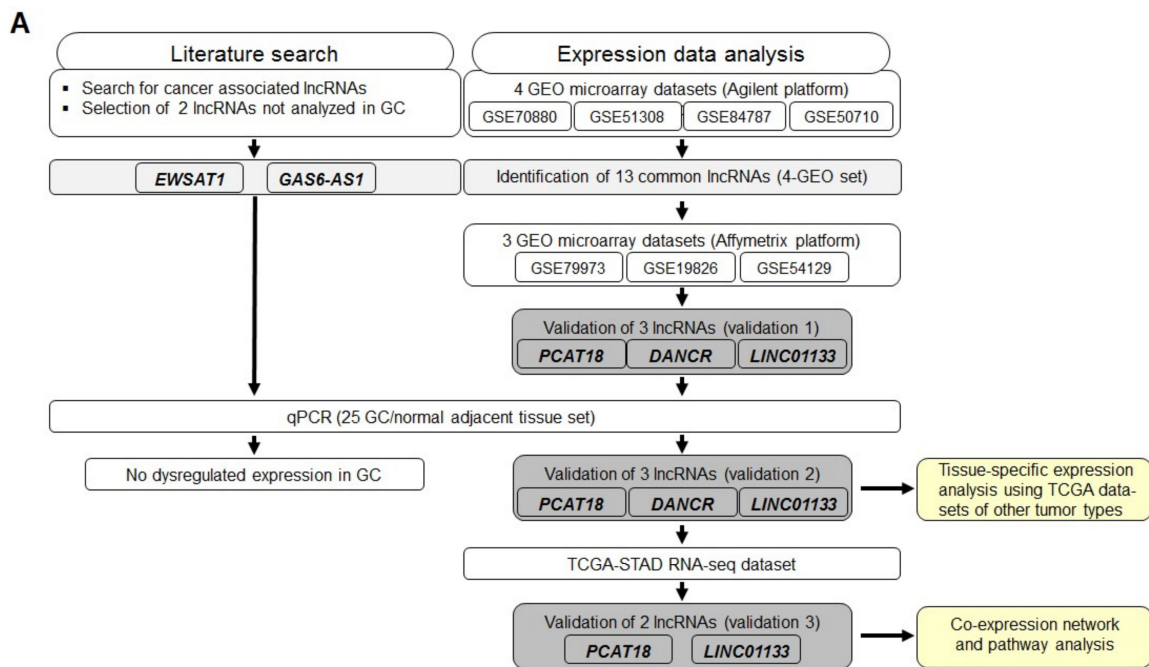
Figure 1. Discovery and selection of long non-coding RNA (lncRNA) candidates. (**A**) Overview of the analysis pipeline; (**B**) Venn diagram showing the number of common and unique lncRNA genes across the GEO datasets. The blue circle represents the lncRNAs found in GSE79973, the pink circle those found in GSE19826, the green circle those found in GSE54129, and the yellow circle the 13 lncRNAs common to the four-GEO set (four Agilent microarray expression datasets of gastric cancer (GC)).
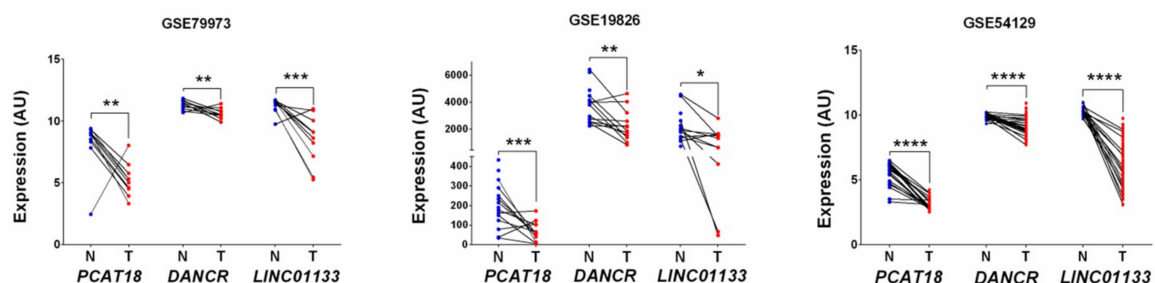

Figure 2. *PCAT18*, *DANCR*, and *LINC01133* expression levels in GC and adjacent normal tissues. Expression data were retrieved from the GEO datasets (**A**) GSE79973, (**B**) GSE19826, and (**C**) GSE54129. N, normal tissue (blue dots); T, tumor tissue (red dots); AU, arbitrary unit. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

Table 1. Differentially expressed lncRNAs in gastric cancer based on analyses of the four discovery GEO datasets.

Gene ID *	Gene Symbol	Gene Name	Log2FC	*p* Value	Regulatory Status
728606	*PCAT18*	Prostate cancer associated transcript 18	−2.62	1.58×10^{-4}	Down
57291	*DANCR*	Differentiation antagonizing non-protein coding RNA	−0.86	1.37×10^{-3}	Down
283738	*NTRK3-AS1*	Neurotrophic receptor tyrosine kinase 3 antisense RNA 1	1.24	2.0×10^{-4}	Up
101928223	*LINC01612*	Long intergenic non-protein coding RNA 1612	−1.50	2.4×10^{-4}	Down
404201	*WDFY3-AS2*	WD repeat and FYVE domain containing 3 antisense RNA 2	−0.77	8.85×10^{-4}	Down
100505875	*LINC01088*	Long intergenic non-protein coding RNA 1088	−2.69	4.44×10^{-9}	Down
102723487	*LINC01497*	Long intergenic non-protein coding RNA 1497	−1.42	1.82×10^{-7}	Down
100505633	*LINC01133*	Long intergenic non-protein coding RNA 1133	−1.42	9.32×10^{-3}	Down
348761	*SPATA3-AS1*	Spermatogenesis associated 3 antisense RNA 1	−1.96	2.02×10^{-9}	Down
401172	*LINC02145*	Long intergenic non-protein coding RNA 2145	−1.95	3.79×10^{-9}	Down
100507099	*FRY-AS1*	FRY microtubule binding protein antisense RNA 1	−1.30	0.02946	Down
286190	*LACTB2-AS1*	Lactamase beta 2 antisense RNA 1	−1.30	0.02946	Down
101927171	*LINC01564*	Long intergenic non-protein coding RNA 1564	−3.30	0.03868	Down

Log2FC < 0: down-regulated, Log2FC > 0: up-regulated. * From NCBI RefSeqGene.

2.2. LncRNA Expression Levels in the 25 GC/Normal Tissue Sample Set and Associations with Clinical, Histopathological, and Epidemiological Parameters

Next we analyzed expression of the five lncRNAs candidates in the 25 GC/normal tissue sample set using real-time quantitative (q) PCR. Expression data were validated for *PCAT18*, *DANCR*, and *LINC01133* (validation 2). All three lncRNAs were down-regulated in GC tissues (*PCAT18*: $p \leq 0.0001$; *DANCR*: $p \leq 0.001$; and *LINC01133*: $p \leq 0.001$) (Figure 3A). No dysregulated expression was observed for *EWSAT1* and *GAS6-AS1* ($p = 0.1034$ and $p = 0.3049$, respectively) and therefore these two lncRNAs were excluded from any further analyses.
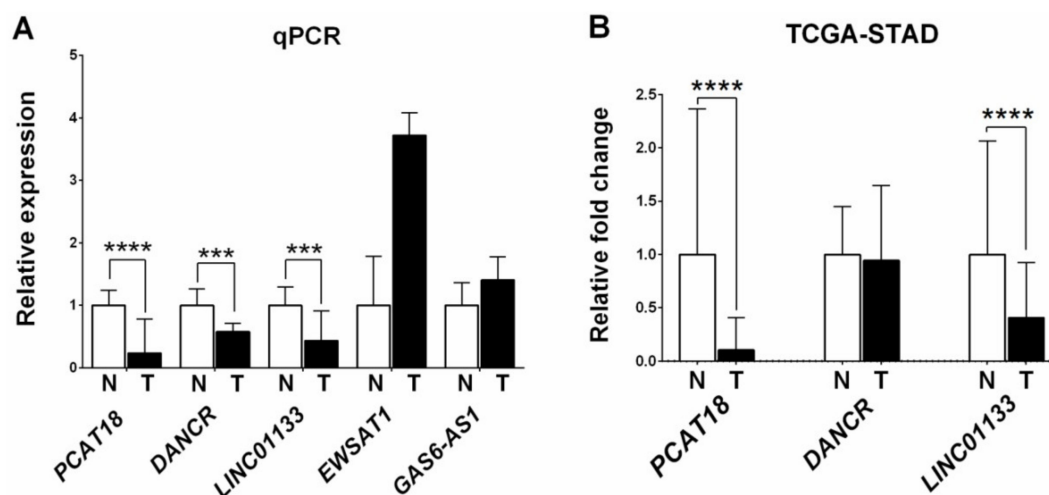

Figure 3. Expression levels of *PCAT18*, *DANCR*, *LINC01133*, *EWSAT1*, and *GAS6-AS1* in GCs and their adjacent normal tissues. (**A**) Expression was measured in 25 GCs and their adjacent normal tissues by real-time qPCR; (**B**) Expression data were retrieved from the TCGA-STAD database. *** $p \leq 0.001$, **** $p \leq 0.0001$.

Table 2 shows the distribution of the 25 GCs according to selected clinical, histopathological, and epidemiological parameters and associations with *PCAT18*, *DANCR*, and *LINC01133* expression levels. With the exception of gender and smoking status, parameters showed no associations with expression levels. *DANCR* expression was associated with gender ($p = 0006$) and smoking status ($p = 0.005$), whereas *LINC01133* expression was associated with gender ($p = 0.002$). Both *LINC01133* and *DANCR* lncRNA expression levels were higher in females than males. *DANCR* lncRNA expression was higher in non-smokers compared to smokers. *PCAT18* expression was not associated with any of the investigated parameters.

Table 2. Associations of clinical and histopathological tumor parameters with *PCAT18*, *DANCR*, and *LINC01133* expression levels.

Parameter	Cases *n* (%)	*p* Value		
		PCAT18	*DANCR*	*LINC01133*
Age (years)				
<60	12 (48)	0.623	0.965	0.504
≥60	13 (52)			
Gender				
Male	21 (84)	0.856	**0.006**	**0.002**
Female	4 (16)			
Site of primary tumor				
Gastric cardia	4 (16)			
Antrum	5 (20)	0.758	0.359	0.536
Stomach	16 (64)			
Tumor size (cm)				
<5	11 (44)			
≥5	12 (48)	0.597	0.809	0.666
Unknown	2 (8)			
Histological grade				
1, 2	17 (68)	0.316	0.879	0.982
3, 4	8 (32)			
Lymph node status				
N0	19 (76)	0.404	0.831	0.771
≥N1	6 (24)			
Vascular invasion				
Yes	19 (76)	0.404	0.831	0.771
No	6 (24)			
Perineural invasion				
Yes	16 (64)	0.966	0.843	0.938
No	9 (36)			
Serosal invasion				
Absent	11 (44)	0.243	0.913	0.732
Present	14 (56)			
Clinical stage				
1, 2	9 (36)	0.223	0.374	0.221
3, 4	16 (64)			
Family history of gastric cancer				
Yes	9 (36)	0.798	0.663	0.431
No	16 (64)			
Smoking status				
Ever	14 (56)	0.056	**0.005**	0.057
Never	11 (44)			

Significant *p* values are marked in bold.

2.3. Analysis of LncRNA Expression Levels Using RNA-Sequencing Data from The Cancer Genome Atlas-Stomach Adenocarcinoma (TCGA-STAD)

For further validation TCGA RNA-seq expression data for *PCAT18*, *DANCR*, and *LINC01133* in stomach cancer and normal tissues were analyzed. As shown in Figure 3B, the expression of *PCAT18* and *LINC01133* was down-regulated in stomach cancer tissues ($p = 6.35 \times 10^{-19}$ and $p = 0.3.40 \times 10^{-7}$, respectively; validation 3). No deregulated expression was observed for *DANCR*.

The relationship between expression of *PCAT18* and *LINC01133*, overall survival and tumor grade/stage was investigated using the The Cancer Genome Atlas- Stomach Adenocarcinoma (TCGA-STAD) dataset. No statistically significant associations were identified.

Further analysis of RNA-seq expression data of these genes in six other cancer entities including four gastrointestinal cancers (esophageal, colon, rectum, and liver), prostate, and breast cancers showed an up-regulation of *PCAT18* in prostate and breast cancers ($p = 2.96 \times 10^{-11}$ and $p = 6.33 \times 10^{-8}$, respectively), and a down-regulation in all gastrointestinal cancers (Figure 4). *LINC01133* was also down-regulated in esophageal, colon, and rectal cancers as well as in prostate cancer. In contrast, *DANCR* was up-regulated in breast, prostate, liver, and colon cancers.
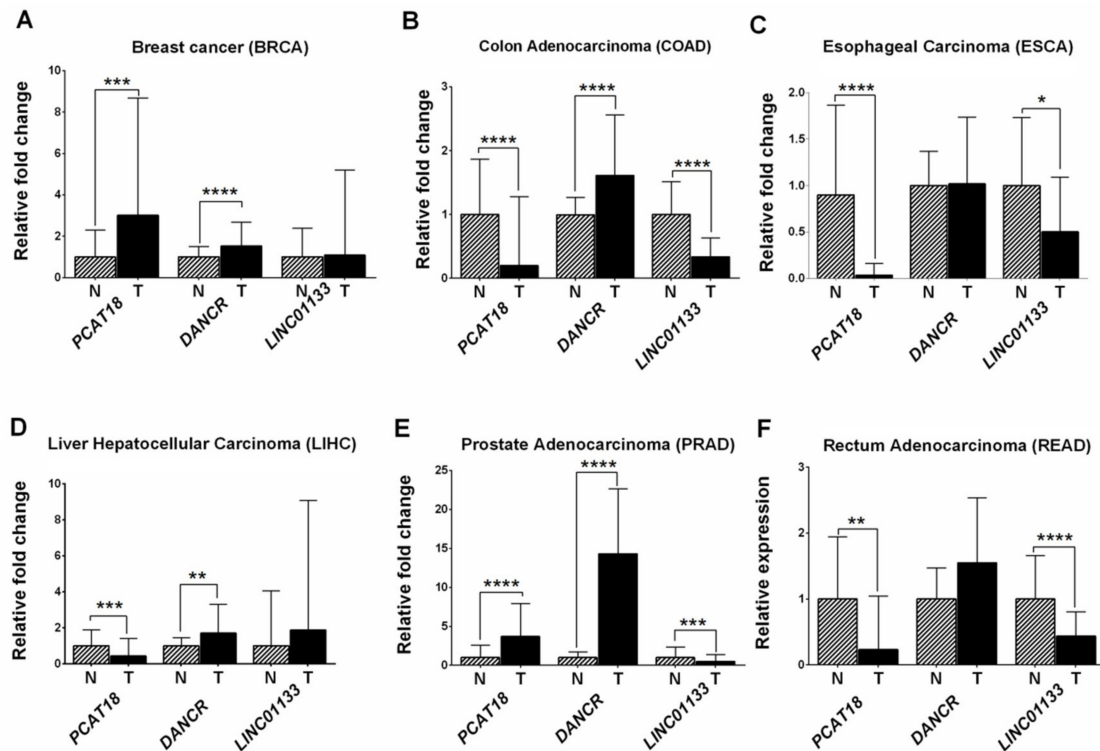

Figure 4. Expression levels of *PCAT18*, *DANCR*, and *LINC01133* in various tumor entities and their adjacent normal tissues. (**A**) Breast cancer (BRCA); (**B**) colon adenocarcinoma (COAD); (**C**) esophageal carcinoma (ESCA); (**D**) liver hepatocellular carcinoma (LIHC); (**E**) prostate adenocarcinoma (PRAD); and (**F**) rectum adenocarcinoma (READ). Expression data were retrieved from TCGA. * $p \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$, **** $p \leq 0.0001$.

2.4. Weighted Gene Correlation Network Analysis and Functional Annotation of Co-Expressed Genes

WGCNA identified 17 co-expression clusters/modules of co-expressed genes in the TCGA-STAD RNA-seq dataset. The two down-regulated lncRNAs *PCAT18* and *LINC01133* were clustered in the red module (Figure 5A). This module contained 218 other genes which are provided in Table S3. The relation between the co-expressed genes of the red module and the lncRNAs is shown in Figure 5B.

To assess how lncRNA genes may contribute to GC development, functional enrichment analysis of the 218 co-expressed genes was performed using Ingenuity Pathway Analysis (IPA). The top ten canonical pathways and the top ten diseases and functions related to the co-expressed genes are shown in Figure 5C,D. The top canonical pathways included different metabolic pathways which play a role in degradation of glycolysis side-product, amino acid biosynthesis, and retinoate biosynthesis. The biological functions of these genes were associated with molecular transport, organismal development, digestive system development and function, organ morphology, and tissue morphology.
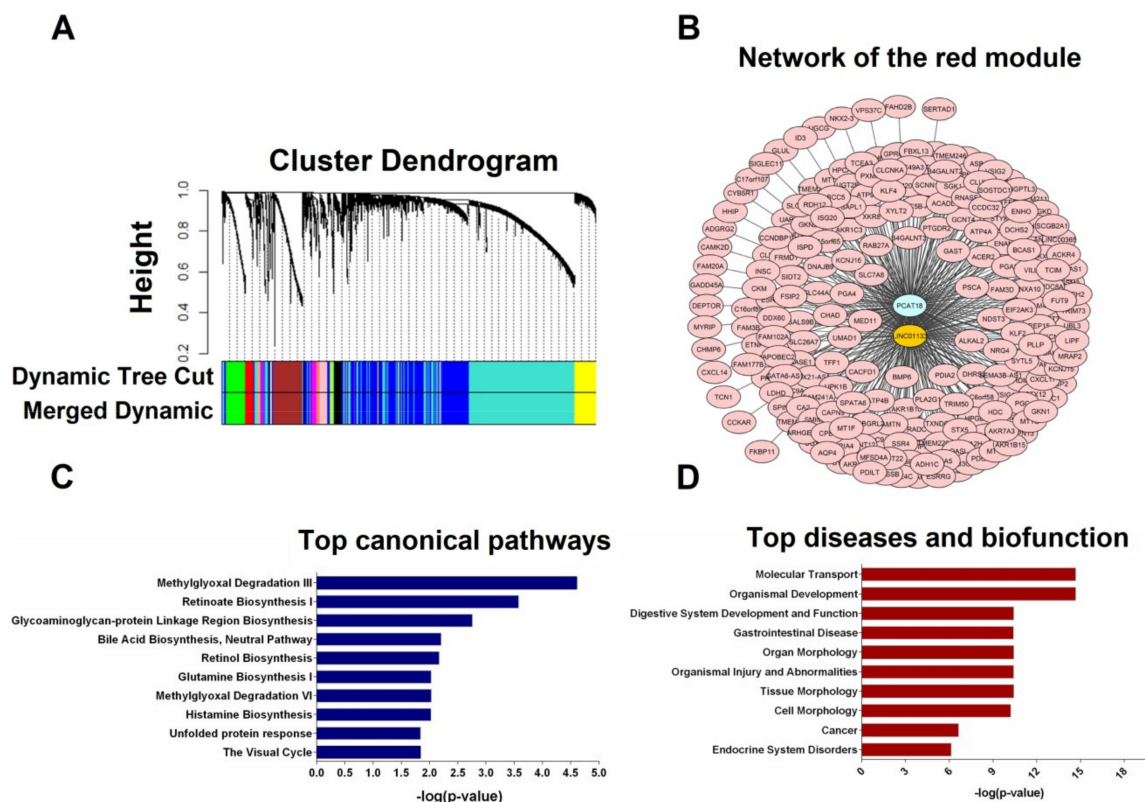
Figure 5. Weighted gene correlation network analysis (WGCNA) and pathway analysis using RNA-sequencing (RNA-seq) data from the TCGA stomach cancer dataset. (**A**) WGCNA cluster dendrogram of differentially expressed mRNAs and lncRNAs. Each leaf (vertical line) in the dendrogram corresponds to a gene. The color row underneath the dendrogram shows the assigned original module and the merged module. The lncRNA candidates *PCAT18* and *LINC01133* were clustered in the red module; (**B**) The lncRNA-mRNA network of the red module visualized using the cytoscape software platform; (**C**) Over-represented terms in the Molecular and Cellular Function category; (**D**) Top diseases and biofunctions related to co-expressed genes.

3. Discussion

Despite advances in diagnosis, prognosis, and treatment, GC remains a worldwide public health concern. Our contribution to this area of investigation was to identify novel lncRNAs involved in gastric carcinogenesis by analyzing gene expression datasets obtained from GC and normal tissues from the GEO and TCGA databases in addition to another sample set of 25 GC and paired normal tissues. While some lncRNAs dysregulated in GC and their clinical value as potential biomarkers for diagnosis and prognosis have been previously reported [22], this study provides additional data on two novel lncRNAs contributing to GC and their potential functions.

LncRNAs are non-coding transcripts longer than 200 nucleotides that do not overlap with annotated coding genes. These transcripts are involved in chromatin remodeling and genome architecture, RNA stabilization and transcriptional regulation [23]. In the present study two lncRNAs down-regulated in GC compared to normal tissues were identified; *PCAT18* and *LINC01133*. The expression data of the lncRNAs are robust, as they were validated in three independent sample sets. This data implies that these lncRNAs may act as tumor suppressors acting during the development of GC.

The prostate cancer associated transcript 18 (*PCAT18*) gene located at 18q11.2 is highly expressed in prostate cancer. Its' silencing in prostate cancer cells leads to the inhibition of cell proliferation, migration, and invasion [24]. In the present study we observed a down-regulation of *PCAT18* in GC/stomach cancer tissues pointing to a role of this lncRNA in GC development. A down-regulation

was also observed in other gastrointestinal cancers including cancers of the esophagus, colon, rectum, and liver, while an up-regulation was observed in breast and prostate cancers. These findings imply that *PCAT18* down-regulation is specific for gastrointestinal tumors. Furthermore, based on normal tissue RNA-seq expression data from the Genotype-Tissue Expression database (GTEx), *PCAT18* is highly expressed in normal stomach tissue [25]. Its high expression suggests a potential regulatory function in this tissue and, when down-regulated, may be a cause or consequence of stomach cancer development.

The long intergenic non-coding RNA 1133 (*LINC01133*) gene on chromosome 1q23.2 is down-regulated in colorectal cancer [26], while an up-regulation was reported in different types of lung cancer [27,28]. In non-small cell lung cancer (NSCLC), its expression inversely correlated with the expression of *KLF*, *P21*, and E-cadherin suggesting an oncogenic function in NSCLC. Furthermore, LINC01133 was shown to sponge the miR-422a to aggravate the tumorigenesis of human osteosarcoma [29]. Our data of a role for *LINC01133* in GC are in line with those from a recent study that showed an association of a reduced *LINC01133* expression with aggressive tumor phenotypes [30]. The authors also showed that this lncRNA inhibits GC progression and metastasis implying its potential use as an anti-metastatic therapeutic target for this disease. Similar to *PCAT18*, *LINC01133* is also highly expressed in normal stomach tissues based on GTEx data implying that its deregulation in normal stomach tissue may play a role in the fate of cells and cancer progression.

Differentiation antagonizing non-protein coding RNA (*DANCR*) located at 4q12 is up-regulated in various cancers including hepatocellular carcinoma [31,32], colorectal cancer [33], prostate cancer [34], osteosarcoma [35], and stomach cancer [36]. In contrast to the previous stomach cancer study, a down-regulation of *DANCR* in GC was found in the present study using the GEO and qPCR data, which however was not validated in the TCGA dataset. Thus, further studies on the expression of this lncRNA in GC and its function are warranted.

The gene co-expression network analysis using WGCNA was performed to identify modules containing *PCAT18*, *LINC01133*, and their co-expressed genes. Seventeen modules were identified, one of which was containing both lncRNAs and 218 eigengenes. Pathway analysis revealed that the top canonical pathways were mostly related to various cell metabolic pathways. Given that tumor cells often have an altered metabolism to cope with the demand of cell-mass increase during growth [37], these lncRNAs may be involved in the control of some metabolic pathways in GC cells. Among the networks, the top ones were linked to molecular transport, organismal development, and gastrointestinal disease. The functional annotations of these top networks were associated with various biofunctions and diseases of the stomach. Merging of the top four networks identified the extracellular-signal-regulated kinase/mitogen-activated protein kinases (ERK/MAPK) pathway as a hub with more connections to other co-expressed genes. It was reported that abnormal activation and mutations of genes involved in the ERK/MAPK pathway occur in more than 50% of human cancer types [38]. Recently, various studies have shown that the ERK/MAPK pathway is involved in regulating cellular mobility in GC cell lines suggesting that this pathway influences GC cell migration and invasion [39]. Besides, of the top upstream regulators, the homeobox genes *CDX1* and *CDX2* have been reported to be crucial players in stomach carcinogenesis [40], while *XBP1* was shown to control the maturation of gastric zymogenic cells [41].

To elucidate the mechanisms of how these lncRNAs exert their functions, lncRNA–lncRNA interactions were predicted in silico using lncRNA2-target databases. An interaction between PCAT18 and BANCR was predicted. *PCAT18* expression was increased following knock-down of *BANCR* [42]. Interestingly, a recent study on GC reported that *BANCR* was significantly up-regulated in GC tissues, and cell lines and its down-regulation led to the inhibition of GC cell proliferation [43]. Accordingly, down-regulation of *PCAT18* along with up-regulation of *BANCR* in GC tissues suggests a possible regulatory interaction between these two lncRNAs. Further exploration using AnnoLnc [44], a web server, which provides systematic annotation of newly identified human lncRNAs, predicted an interaction of PCAT18 with C15orf57 (now called *CCDC32* gene) and GABRR3 protein.

Our study has some limitations that are related to the small size of the internal validation set, and to the retrospective study cohorts. Another limitation is the lack of functional analyses, which should be performed to yield detailed insight into the mechanism of downregulation of *PCAT18* and *LINC01133* in gastric carcinogenesis.

Altogether, we showed a decreased tissue-specific expression of *PCAT18* and *LINC01133* lncRNAs in GC and other gastrointestinal tumor tissues, suggesting a role of these lncRNAs in the development of gastrointestinal tumors. The reduced lncRNAs expression levels may interfere with normal harmony of gene regulation in normal gastric cells and potentiate them towards GC progression and development, which may be achieved via a gene regulation process leading to metabolic adaptation in tumor cells. The two lncRNA biomarkers may contribute to a better understanding of the complex mechanisms of gastric carcinogenesis. The reported data should guide future studies on the associations of *PCAT18* and *LINC01133* with GC and their functions.

4. Materials and Methods

4.1. Data Extraction from the GEO Database and Literature Review

Ten GC datasets were retrieved from GEO (http://www.ncbi.nlm.nih.gov/geo/) using the keywords: "lncRNA stomach cancer" (study keyword), "homo sapiens" (organism), "expression profiling by array" (study type), and "tissue" (attribute name). Seven datasets fulfilling the following parameters were selected for expression analyses: (1) Availability of data on GC and adjacent normal tissues; (2) inclusion of expression data of lncRNA genes; and (3) availability of minimum information about the microarray experiment. Four datasets (obtained using the Agilent platform);, GSE70880 (20 tumor and 20 adjacent normal tissues), GSE51308 (5 tumor and 5 adjacent normal tissues), GSE84787 (10 tumor and 10 adjacent normal tissues), and GSE50710 (10 tumor and 10 adjacent normal tissues) were selected for the discovery of lncRNA candidates. These sets contained data from 45 GCs and their paired normal tissues. Besides, three datasets (obtained using the Affymetrix platform); GSE79973 (10 tumor and 10 adjacent normal tissues), GSE19826 (12 tumor and adjacent normal tissues + 3 normal gastric tissues), and GSE54129 (111 tumor and 21 noncancerous gastric tissues), were used for data validation. These sets contained data from 133 GCs and 46 noncancerous normal tissues.

The comparison between tumor and adjacent normal tissues allowed the identification of differentially expressed genes in the GEO datasets. p values were adjusted (p_{adj}.) using the Benjamini and Hochberg method. A p_{adj}. < 0.05 and a $|logFC| \geq 1$ were set as cut-off criteria [45–47]. Among the top candidates, those already known to be associated with GC were excluded.

Moreover, in order to identify novel lncRNAs associated with other cancers but not reported in GC, a comprehensive PubMed literature search was performed. The following keywords, selected from the medical subject headings (MeSH) database, were used: ("Neoplasms") AND "RNA, Long Noncoding").

4.2. Patient Samples

Fifty tissues comprising 25 GC and paired normal tissues (25 GC/normal tissue sample set) were obtained from the Iran National Tumor Bank (INTB, Tehran, Iran). All tissues were collected during surgical resection of patients diagnosed with primary GC at the Imam Khomeini Hospital, Tehran, Iran from 02/2009 to 11/2014. Adjacent normal tissues were obtained from areas at least 6 cm away from the tumor site. None of the patients received radiation and/or chemotherapy treatment before surgery. Tissues were stored in liquid nitrogen until nucleic acid extraction.

Data on selected clinical, histopathological, and epidemiological parameters including age (<60, ≥60 years), gender, site of primary tumor (gastric cardia, antrum, stomach), tumor size (<5, ≥5 cm), histological grade (1–4), lymph node status (N0, ≥N1), vascular invasion (yes, no), perineural invasion (yes, no) serosal invasion (yes, no), clinical stage (1–4), family history of GC (yes, no), and smoking status (ever, never) are presented in Table 2.

The study was approved by the Ethical Committee of the Shahroud University of Medical Sciences (9559, 08/10/2016). All study participants provided written informed consent.

4.3. RNA Extraction and cDNA Synthesis

Tissues were grinded in liquid nitrogen using a mortar and pestle, instantly transferred into the lysis buffer, and homogenized using a needle and syringe. Total RNA was extracted using AllPrep DNA/RNA Mini kit (Qiagen, Hilden, Germany), according to manufacturer's instructions. The quantity and quality of isolated RNA samples were determined by Picodrop microliter spectrophotometer (OEM, Hinxton, UK), and electrophoresis on a 0.8% agarose gel. Afterwards, 1 µg of total RNA was converted into cDNA using PrimeScript^{TM} RT reagent kit (TaKaRa Bio, Shiga, Japan) according to the manufacturer's instruction.

4.4. Real-Time PCR

Real-Time qPCR was performed using the Bio-Rad CFX96^{TM} Real-Time System (Bio-Rad, Foster City, CA, USA). The 20 µL reaction contained 10 µL SYBER Master Mix without ROX (Ampliqon, 5230 Odense M, Denmark), 8 µL ddH$_2$O, 1 µL cDNA template, 5 µM of each primer (forward and reverse). The real time qPCR primer sequences are provided in the Table S4. Conditions for amplification were 95 °C for 15 min, followed by 40 cycles of 95 °C for 5 s, 60 °C for 30 s. Melting curves were obtained by slow heating (0.5 °C/s) at temperatures in the range of 65 to 95 °C. All samples were run in duplicate.

4.5. Data Extraction from TCGA and Data Analyses

For further data validation, RNA-sequencing data (RNA-seq) of stomach (gastric) adenocarcinoma (STAD) from TCGA were analyzed and used for comparison. Moreover, the expression levels of the lncRNA candidates were analyzed in other cancer entities including adenocarcinomas of the colon, esophagus, rectum, liver hepatocellular, prostate, and breast and adjacent normal tissues. The RNA-seq raw data files were downloaded from the TCGA GDC data portal, normalized, and filtered using the R/Bioconductor software package TCGAbiolinks (R version 3.4.2, http://www.r-project.org/) [48]. Differential expression analysis of lncRNA genes in tumor and normal tissues was performed using the edgeR package [49]. This package implements the trimmed mean of M-values (TMM) method to give the normalized read counts.

4.6. Weighted Gene Correlation Network Analysis

Given that the functions of most lncRNAs are unknown, prediction of their functions mostly relies on the analysis of their co-expressed genes. Network analysis was performed using the WGCNA package in R as described previously [50,51]. To identify modules of highly correlated genes, WGCNA was performed on RNA-seq data of the TCGA-STAD dataset obtained from 407 GC and normal tissues. To identify modules with different expression patterns, a soft threshold power was assigned to create co-expression networks. The networks were built by merging genes with highly similar co-expression patterns into modules and the eigengenes of these modules were determined. Finally, the module with the key lncRNAs and their co-expressed genes was obtained. The reconstructed co-expression network was visualized using the Cytoscape software (version 3.5.1 (http://www.cytoscape.org)) [52].

4.7. Functional Annotation of the Co-Expressed Genes in the Module

To investigate the potential functions of *PCAT18* and *LINC01133* and their associated biological pathways, a functional enrichment analysis of their co-expressed genes was performed using Ingenuity Pathway Analysis (IPA; Ingenuity Systems, Mountain View, CA, USA) software. IPA provides a graphical representation of the molecular relationships between genes in networks. Functional analysis identified statistical significant (Fisher's exact test *p* value < 0.05) over-represented Canonical

Pathways, Molecular and Cellular Functions, Physiological System Development and Function, upstream regulators, and Diseases and Bio Functions in the imported data sets.

4.8. Statistical Analyses

Statistical analyses were performed using GraphPad Prism 6 software (GraphPad Software Inc., San Diego, CA, USA) and SPSS 24.0 (SPSS Inc., Chicago, IL, USA). Differences between the means of two groups were determined using student's *t*-test. All *p* values were two-sided, with a *p* value of less than 0.05 considered statistically significant. All results are presented as the mean ± standard deviation (SD) of the experiments.

5. Conclusions

The down-regulation of *PCAT18* and *LINC01133* in GC implies that these lncRNAs may have a tumor suppressive function in the development of gastric tumors. The two lncRNA biomarkers may contribute to a better understanding of the complex mechanisms of gastric carcinogenesis.

Supplementary Materials: Supplementary materials can be found at http://www.mdpi.com/1422-0067/19/12/3881/s1.

Abbreviations

GC	Gastric Cancer
GEO	Gene Expression Omnibus
TCGA	The Cancer Genome Atlas
lncRNAs	Long Non-Coding RNAs
WGCNA	Weighted Gene Correlation Network Analysis
RNA-seq	RNA-Sequencing
STAD	Stomach Adenocarcinoma
TMM	Trimmed Mean of M-Values
IPA	Ingenuity Pathway Analysis
SD	Standard Deviation
PCAT18	Prostate Cancer Associated Transcript 18
LINC01133	Long Intergenic Non-Coding RNA 1133
DANCR	Differentiation Antagonizing Non-Protein Coding RNA
GTEx	Genotype-Tissue Expression Database
NSCLC	Non-Small Cell Lung Cancer

References

1. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D.M.; Forman, D.; Bray, F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **2015**, *136*, E359–E386. [CrossRef] [PubMed]

2. Bria, E.; de Manzoni, G.; Beghelli, S.; Tomezzoli, A.; Barbi, S.; di Gregorio, C.; Scardoni, M.; Amato, E.; Frizziero, M.; Sperduti, I.; et al. A clinical-biological risk stratification model for resected gastric cancer: Prognostic impact of Her2, Fhit, and APC expression status. *Ann. Oncol.* **2013**, *24*, 693–701. [CrossRef] [PubMed]

3. Satolli, M.A.; Buffoni, L.; Spadi, R.; Roato, I. Gastric cancer: The times they are a-changin. *World J. Gastrointest. Oncol.* **2015**, *7*, 303–316. [CrossRef]

4. Shimizu, T.; Marusawa, H.; Matsumoto, Y.; Inuzuka, T.; Ikeda, A.; Fujii, Y.; Minamiguchi, S.; Miyamoto, S.; Kou, T.; Sakai, Y.; et al. Accumulation of somatic mutations in TP53 in gastric epithelium with Helicobacter pylori infection. *Gastroenterology* **2014**, *147*, 407-17.e3. [CrossRef] [PubMed]

5. Nagini, S. Carcinoma of the stomach: A review of epidemiology, pathogenesis, molecular genetics and chemoprevention. *World J. Gastrointest. Oncol.* **2012**, *4*, 156–169. [CrossRef] [PubMed]

6. Gigek, C.O.; Chen, E.S.; Calcagno, D.Q.; Wisnieski, F.; Burbano, R.R.; Smith, M.A. Epigenetic mechanisms in gastric cancer. *Epigenomics* **2012**, *4*, 279–294. [CrossRef]

7. Bartonicek, N.; Maag, J.L.V.; Dinger, M.E. Long noncoding RNAs in cancer: Mechanisms of action and technological advancements. *Mol. Cancer* **2016**, *15*, 43. [CrossRef]

8. Sun, W.; Yang, Y.; Xu, C.; Guo, J. Regulatory mechanisms of long noncoding RNAs on gene expression in cancers. *Cancer Genet.* **2017**, *216–217*, 105–110. [CrossRef] [PubMed]

9. Wang, K.C.; Chang, H.Y. Molecular mechanisms of long noncoding RNAs. *Mol. Cell* **2011**, *43*, 904–914. [CrossRef]

10. Böhmdorfer, G.; Wierzbicki, A.T. Control of chromatin structure by long noncoding RNA. *Trends Cell Biol.* **2015**, *25*, 623–632. [CrossRef] [PubMed]

11. Schmitt, A.M.; Chang, H.Y. Long noncoding RNAs in cancer pathways. *Cancer Cell* **2016**, *29*, 452–463. [CrossRef] [PubMed]

12. Yan, X.; Hu, Z.; Feng, Y.; Hu, X.; Yuan, J.; Zhao, S.D.; Zhang, Y.; Yang, L.; Shan, W.; He, Q.; et al. Comprehensive genomic characterization of long non-coding RNAs across human cancers. *Cancer Cell* **2015**, *28*, 529–540. [CrossRef] [PubMed]

13. Brunner, A.L.; Beck, A.H.; Edris, B.; Sweeney, R.T.; Zhu, S.X.; Li, R.; Montgomery, K.; Varma, S.; Gilks, T.; Guo, X.; et al. Transcriptional profiling of long non-coding RNAs and novel transcribed regions across a diverse panel of archived human cancers. *Genome Biol.* **2012**, *13*, R75. [CrossRef] [PubMed]

14. Zhang, Y.; Xu, Y.; Feng, L.; Li, F.; Sun, Z.; Wu, T.; Shi, X.; Li, J.; Li, X. Comprehensive characterization of lncRNA-mRNA related ceRNA network across 12 major cancers. *Oncotarget* **2016**, *7*, 64148–64167. [CrossRef] [PubMed]

15. Lin, C.; Yang, L. Long noncoding RNA in cancer: Wiring signaling circuitry. *Trends Cell Biol.* **2018**, *28*, 287–301. [CrossRef] [PubMed]

16. Fang, X.Y.; Pan, H.F.; Leng, R.X.; Ye, D.Q. Long noncoding RNAs: Novel insights into gastric cancer. *Cancer Lett.* **2015**, *356 (2 Pt B)*, 357–366. [CrossRef]

17. Yang, F.; Bi, J.; Xue, X.; Zheng, L.; Zhi, K.; Hua, J.; Fang, G. Up-regulated long non-coding RNA H19 contributes to proliferation of gastric cancer cells. *FEBS J.* **2012**, *279*, 3159–3165. [CrossRef] [PubMed]

18. Endo, H.; Shiroki, T.; Nakagawa, T.; Yokoyama, M.; Tamai, K.; Yamanami, H.; Fujiya, T.; Sato, I.; Yamaguchi, K.; Tanaka, N.; et al. Enhanced expression of long non-coding RNA HOTAIR is associated with the development of gastric cancer. *PLoS ONE* **2013**, *8*, e77070. [CrossRef]

19. Li, L.; Zhang, L.; Zhang, Y.; Zhou, F. Increased expression of LncRNA BANCR is associated with clinical progression and poor prognosis in gastric cancer. *Biomed. Pharmacother.* **2015**, *72*, 109–112. [CrossRef] [PubMed]

20. Sun, M.; Xia, R.; Jin, F.; Xu, T.; Liu, Z.; De, W.; Liu, X. Downregulated long noncoding RNA MEG3 is associated with poor prognosis and promotes cell proliferation in gastric cancer. *Tumour Biol.* **2014**, *35*, 1065–1073. [CrossRef]

21. Sun, M.; Jin, F.Y.; Xia, R.; Kong, R.; Li, J.H.; Xu, T.P.; Liu, Y.W.; Zhang, E.B.; Liu, X.H.; De, W. Decreased expression of long noncoding RNA GAS5 indicates a poor prognosis and promotes cell proliferation in gastric cancer. *BMC Cancer* **2014**, *14*, 319. [CrossRef] [PubMed]

22. Yang, Z.; Guo, X.; Li, G.; Shi, Y.; Li, L. Long noncoding RNAs as potential biomarkers in gastric cancer: Opportunities and challenges. *Cancer Lett.* **2016**, *371*, 62–70. [CrossRef] [PubMed]

23. Ransohoff, J.D.; Wei, Y.; Khavari, P.A. The functions and unique features of long intergenic non-coding RNA. *Nat. Rev. Mol. Cell Biol.* **2018**, *19*, 143–157. [CrossRef]

24. Crea, F.; Watahiki, A.; Quagliata, L.; Xue, H.; Pikor, L.; Parolia, A.; Wang, Y.; Lin, D.; Lam, W.L.; Farrar, W.L.; et al. Identification of a long non-coding RNA as a novel biomarker and potential therapeutic target for metastatic prostate cancer. *Oncotarget* **2014**, *5*, 764–774. [CrossRef] [PubMed]

25. GTEx Consortium. Genetic effects on gene expression across human tissues. *Nature* **2017**, *550*, 204–213. [CrossRef]

26. Zhang, J.H.; Li, A.Y.; Wei, N. Downregulation of long non-coding RNA LINC01133 is predictive of poor prognosis in colorectal cancer patients. *European Rev. Med. Pharmacol. Sci.* **2017**, *21*, 2103–2107.

27. Zhang, J.; Zhu, N.; Chen, X. A novel long noncoding RNA LINC01133 is upregulated in lung squamous cell cancer and predicts survival. *Tumour Biol.* **2015**, *36*, 7465–7471. [CrossRef]

28. Zang, C.; Nie, F.Q.; Wang, Q.; Sun, M.; Li, W.; He, J.; Zhang, M.; Lu, K.H. Long non-coding RNA LINC01133 represses KLF2, P21 and E-cadherin transcription through binding with EZH2, LSD1 in non small cell lung cancer. *Oncotarget* **2016**, *7*, 11696–11707. [CrossRef]

29. Zeng, H.F.; Qiu, H.Y.; Feng, F.B. Long Noncoding RNA LINC01133 sponges miR-422a to aggravate the tumorigenesis of human osteosarcoma. *Oncol. Res.* **2017**, *26*, 335–343. [CrossRef]

30. Yang, X.Z.; Cheng, T.T.; He, Q.J.; Lei, Z.Y.; Chi, J.; Tang, Z.; Liao, Q.X.; Zhang, H.; Zeng, L.S.; Cui, S.Z. LINC01133 as ceRNA inhibits gastric cancer progression by sponging miR-106a-3p to regulate APC expression and the Wnt/beta-catenin pathway. *Mol. Cancer* **2018**, *17*, 126. [CrossRef]

31. Ma, X.; Wang, X.; Yang, C.; Wang, Z.; Han, B.; Wu, L.; Zhuang, L. DANCR acts as a diagnostic biomarker and promotes tumor growth and metastasis in hepatocellular carcinoma. *Anticancer Res.* **2016**, *36*, 6389–6398. [CrossRef] [PubMed]

32. Yuan, S.X.; Wang, J.; Yang, F.; Tao, Q.F.; Zhang, J.; Wang, L.L.; Yang, Y.; Liu, H.; Wang, Z.G.; Xu, Q.G.; et al. Long noncoding RNA DANCR increases stemness features of hepatocellular carcinoma by derepression of CTNNB1. *Hepatology (Baltimore, Md.)* **2016**, *63*, 499–511. [CrossRef] [PubMed]

33. Liu, Y.; Zhang, M.; Liang, L.; Li, J.; Chen, Y.-X. Over-expression of lncRNA DANCR is associated with advanced tumor progression and poor prognosis in patients with colorectal cancer. *Int. J.Clin. Exp. Pathol.* **2015**, *8*, 11480–11484. [PubMed]

34. Jia, J.; Li, F.; Tang, X.-S.; Xu, S.; Gao, Y.; Shi, Q.; Guo, W.; Wang, X.; He, D.; Guo, P. Long noncoding RNA DANCR promotes invasion of prostate cancer through epigenetically silencing expression of TIMP2/3. *Oncotarget* **2016**, *7*, 37868–37881. [CrossRef] [PubMed]

35. Jiang, N.; Wang, X.; Xie, X.; Liao, Y.; Liu, N.; Liu, J.; Miao, N.; Shen, J.; Peng, T. lncRNA DANCR promotes tumor progression and cancer stemness features in osteosarcoma by upregulating AXL via miR-33a-5p inhibition. *Cancer Lett.* **2017**, *405*, 46–55. [CrossRef] [PubMed]

36. Hao, Y.P.; Qiu, J.H.; Zhang, D.B.; Yu, C.G. Long non-coding RNA DANCR, a prognostic indicator, promotes cell growth and tumorigenicity in gastric cancer. *Tumour Biol.* **2017**, *39*. [CrossRef]

37. Hung, C.L.; Wang, L.Y.; Yu, Y.L.; Chen, H.W.; Srivastava, S.; Petrovics, G.; Kung, H.J. A long noncoding RNA connects c-Myc to tumor metabolism. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 18697–18702. [CrossRef] [PubMed]

38. Dhillon, A.S.; Hagan, S.; Rath, O.; Kolch, W. MAP kinase signalling pathways in cancer. *Oncogene* **2007**, *26*, 3279–3290. [CrossRef]

39. Yang, M.; Huang, C.Z. Mitogen-activated protein kinase signaling pathway and invasion and metastasis of gastric cancer. *World J. Gastroenterol.* **2015**, *21*, 11673–11679. [CrossRef]

40. Joo, M.K.; Park, J.J.; Chun, H.J. Impact of homeobox genes in gastrointestinal cancer. *World J. Gastroenterol.* **2016**, *22*, 8247–8256. [CrossRef]

41. Huh, W.J.; Esen, E.; Geahlen, J.H.; Bredemeyer, A.J.; Lee, A.H.; Shi, G.; Konieczny, S.F.; Glimcher, L.H.; Mills, J.C. XBP1 controls maturation of gastric zymogenic cells by induction of MIST1 and expansion of the rough endoplasmic reticulum. *Gastroenterology* **2010**, *139*, 2038–2049. [CrossRef] [PubMed]

42. Flockhart, R.J.; Webster, D.E.; Qu, K.; Mascarenhas, N.; Kovalski, J.; Kretz, M.; Khavari, P.A. BRAF(V600E) remodels the melanocyte transcriptome and induces BANCR to regulate melanoma cell migration. *Genome Res.* **2012**, *22*, 1006–1014. [CrossRef]

43. Zhang, Z.X.; Liu, Z.Q.; Jiang, B.; Lu, X.Y.; Ning, X.F.; Yuan, C.T.; Wang, A.L. BRAF activated non-coding RNA (BANCR) promoting gastric cancer cells proliferation via regulation of NF-kappaB1. *Biochem. Biophys. Res. Commun.* **2015**, *465*, 225–231. [CrossRef] [PubMed]

44. Hou, M.; Tang, X.; Tian, F.; Shi, F.; Liu, F.; Gao, G. AnnoLnc: A web server for systematically annotating novel human lncRNAs. *BMC Genom.* **2016**, *17*, 931. [CrossRef] [PubMed]

45. Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; Marshall, K.A.; Phillippy, K.H.; Sherman, P.M.; Holko, M.; et al. NCBI GEO: Archive for functional genomics data sets—Update. *Nucleic Acids Res.* **2013**, *41*, D991–D995. [CrossRef] [PubMed]

46. Clough, E.; Barrett, T. The gene expression omnibus database. *Methods Mol. Biol. (Clifton, N.J.)* **2016**, *1418*, 93–110.

47. Benjamini, Y.; Hochberg, Y. Controlling the false discovery Rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. Ser. B (Methodol.)* **1995**, *57*, 289–300.

48. Colaprico, A.; Silva, T.C.; Olsen, C.; Garofano, L.; Cava, C.; Garolini, D.; Sabedot, T.S.; Malta, T.M.; Pagnotta, S.M.; Castiglioni, I.; et al. TCGAbiolinks: An R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res.* **2016**, *44*, e71. [CrossRef]

49. Robinson, M.D.; McCarthy, D.J.; Smyth, G.K. edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **2010**, *26*, 139–140. [CrossRef]

50. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [CrossRef]

51. Zhang, B.; Horvath, S. A general framework for weighted gene co-expression network analysis. *Stat. Appl. Genet. Mol. Biol.* **2005**, *4*. [CrossRef] [PubMed]

52. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [CrossRef] [PubMed]