Original Article
Identification of potential prognostic biomarkers for gastric adenocarcinoma by constructing competing endogenous RNA network

Wei Wang, Zheng Li, Shenyu Wang

Department of Integrated Traditional Chinese and Western Medicine, Cancer Hospital of China Medical University, Liaoning Cancer Hospital and Institute, No. 44 Xiaoheyan Road, Dadong District, Shenyang 110042, Liaoning Province, PR China

Received June 3, 2019; Accepted September 2, 2019; Epub November 15, 2019; Published November 30, 2019

Abstract: Background: Our aim was to predict prognostic biomarkers for gastric adenocarcinoma by constructing competitive endogenous RNA (ceRNA) networks. Methods: The gastric adenocarcinoma data were downloaded from The Cancer Genome Atlas (TCGA) database. Differentially expressed long non-coding RNA (lncRNA), microRNAs (miRNAs), and mRNAs between gastric adenocarcinoma tissue samples and normal samples were selected using edger package. After matching the lncRNA-miRNA and miRNA-mRNA relationships, the ceRNA network was constructed using Cytoscape v3.6. After calculating the degree of lncRNAs in the ceRNA network, we constructed the sub-network. After that, the functional enrichment analysis of mRNAs in the sub-network was performed by use of Cytoscape plug-in BinGO. Finally, we performed a survival analysis to evaluate the prognostic performance of potential biomarkers. Results: In total, 1635 mRNAs, 1002 lncRNAs, and 98 miRNAs were differentially expressed in gastric adenocarcinoma samples compared with normal tissue samples. After matching, there were 149 lncRNA-miRNA relationships and 15 miRNA-mRNA relationships. Then we constructed the ceRNA network. After calculating the degree of lncRNAs in the ceRNA network, 2 lncRNA with degree > 7 were screened to construct the sub-network (2 lncRNAs, 11 mRNAs and 8 miRNAs). Functional enrichment analysis revealed that the mRNAs in the sub-network were involved in a variety of biological processes such as negative regulation of cell adhesion mediated by integrin. According to the survival analysis, SERPINE1 and ADAMTS9-AS2 were closely related with patients’ survival time. Conclusion: We found that SERPINE1 and ADAMTS9-AS2 can become potential prognostic biomarkers for gastric adenocarcinoma patients.

Keywords: Gastric adenocarcinoma, competing endogenous RNA network, sub-network, SERPINE1, ADAMTS9-AS2

Introduction

Gastric cancer is the second leading cause of cancer-related death worldwide, and is the fourth most common malignancy in China [1]. Gastric adenocarcinoma is a type of gastric cancer, 95% of which develop into malignant tumors. Although the treatment of cancer is constantly updated, the current clinical strategy is still very backward in determining the survival prognosis of gastric adenocarcinoma patients due to the limited understanding of the pathogenesis [2]. Therefore, exploring new molecular targets can help predict the prognosis of gastric adenocarcinoma and the development of prognosis-targeted drugs. In recent years, many reports have begun to try to explore the mechanism of gastric adenocarcinoma from a new perspective-competitive endogenous RNA. Competitive endogenous RNAs (ceRNAs) are small molecules that modulate transcription at the post-transcriptional level through binding to miRNAs competitively. In 2007, Franco Zorrilla and others reported a discovery, which they called “target mimicry” [3]. In 2011, the word “CERNA” was used to describe new molecules that regulate post-transcriptional regulation [4]. MicroRNAs (miRNAs) are a class of endogenous single-stranded, non-coding, small RNAs that regulate gene expression at the transcriptional or post-transcriptional level in eukaryotes through sequen-
Potential prognostic biomarkers for gastric adenocarcinoma

Competitive endogenous RNAs (ceRNAs) regulate gene expression by miRNA response components (MREs) binding to miRNAs [6]. With the exception of approximately 2% of protein-coding genes, the vast majority of the human genome consists of non-coding RNAs (ncRNAs) that are dispersed throughout the genomic DNA. A subcategory of ncRNAs is called long non-coding RNA (lncRNA) which contains a length of more than 200 nucleotides [7]. LncRNA can be used as a ceRNA to play a role in many different cancer biological pathways, including proliferation, apoptosis, adhesion, angiogenesis and metastasis [8]. It has been reported that lncRNAs play an important role in the occurrence and metastasis of gastric adenocarcinoma [9]. At present, several lncRNAs related to the occurrence and metastasis of gastric cancer have been found; such as GHET1, H19, CCAT2, HNF1A-AS1, FENDRR and so on [10-13]. However, unlike the protein-coding genes, most lncRNA sequences are poorly conserved and cannot be directly inferred by the base sequence [14].

Therefore, in this study, the ceRNA network was constructed to explore the significantly expressed lncRNAs in gastric adenocarcinoma, and the lncRNA-miRNA-mRNA sub-network was constructed. We performed a survival analysis to find out the potential prognostic molecular markers for gastric adenocarcinoma.

Table 1. Fifteen miRNA-mRNA relationships related with gastric adenocarcinoma

<table>
<thead>
<tr>
<th>miRNA</th>
<th>mRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-mir-145</td>
<td>MEST, SERPINE1</td>
</tr>
<tr>
<td>hsa-mir-204</td>
<td>CHRDL1, HOXC8, IL11, NPTX1</td>
</tr>
<tr>
<td>hsa-mir-205</td>
<td>ESRRG</td>
</tr>
<tr>
<td>hsa-mir-372</td>
<td>ATAD2, CADM2, LEFTY1, TMEM100</td>
</tr>
<tr>
<td>hsa-mir-373</td>
<td>ATAD2, CADM2, LEFTY1, TMEM100</td>
</tr>
</tbody>
</table>

Figure 1. Differential expression analysis. A-C. Volcano plots of mRNAs, lncRNAs and miRNAs. D-F. Heat maps of mRNAs, lncRNAs and miRNAs. Red represents up-regulated genes and green stands for down-regulated genes.
Potential prognostic biomarkers for gastric adenocarcinoma

**Materials and methods**

**Samples and preprocessing**

The transcriptional data of the STAD-project (Stomach Adenocarcinoma) were downloaded from the cancer genome atlas (TCGA) database, including the count matrix of RNA expression and miRNA-seq data [15].

This table shows that 43 up-regulated lncRNAs and 20 down-regulated lncRNAs are in the ceRNA network.

<table>
<thead>
<tr>
<th>Expression</th>
<th>IncRNAs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Down-regulated lncRNAs</td>
<td>ADAMTS9-AS1, HCG22, AC011374.1, LRRC3-AS1, AC024597.1, ADAMTS9-AS2, AC092422.1, RBMS3-AS3, LINC00330, ARHGGEF26-AS1, AC110491.1, C20orf166-AS1, PART1, IL20RB-AS1 AL357153.1, FRMD6-AS2, LINC00332, LINC00163, LINC00284, MIR205HG, TDRG1.</td>
</tr>
<tr>
<td>Up-regulated lncRNAs</td>
<td>HOTAIR, AP002478.1, C7orf69, C8orf31, C20orf48, LINC00355, C15orf54, HOTTIP, DLX6-AS1, PLCH1-AS2, LINC00501, HULC, UCA1, ERVMER61-1, POU6F2-AS2, VCAN-AS1, LINC00184, AL391832.1, AN01-AS2, DSCR4-IT1, AC034229.1, LINC00534, AC006449.1, LINC00524, IGF2-AS, LINC0052, AC061975.6, LINC00326, BOK-AS1, LINC00114, CECR3, LINC00393, AP003027.1, AC090398.1, LINC00200, C17orf77, LINC00410, AC026320.1, LINC00221, NKX2-1-AS1, LINC00523, AC010145.1.</td>
</tr>
</tbody>
</table>

This table shows that 43 up-regulated lncRNAs and 20 down-regulated lncRNAs are in the ceRNA network.
matrix. According to the sample barcode, the samples were divided into cancer samples and normal samples for further analysis.

**Differential expression analysis**

Differential expression analysis was performed using the edger package and results were visualized using heatmap package. The screening threshold of differentially expressed IncRNAs, miRNAs and mRNAs was as follows: logFC absolute value > 2 and FDR (adjusted P value) < 0.05.

The ceRNA relationship match

The ceRNA relationship matching was divided into three steps: (i) the obtained differential expressed IncRNAs and miRNAs were extracted and matched with the known pairs of IncRNA-miRNAs in the IncRNA and miRNA target prediction database “miRcode” (http://mirdb.org) to obtain IncRNA-miRNA related pairs related to gastric adenocarcinoma [16]; (ii) the database “starBase” was used to add differentially expressed miRNAs with 3p or 5p anchors; (iii) differentially expressed miRNAs and mRNAs were extracted, and the corresponding miRNA-mRNA relationship pair was matched in the miRNA target gene database “miRDB”, “miRTarBase” (http://miRTarbase.mbc.nctu.edu.tw/index.html), “TargetScan” (http://www.targetscan.org/) to obtain the miRNA-mRNA relationships related to gastric adenocarcinoma.

The construction of the ceRNA network

The IncRNA-miRNA and miRNA-mRNA relationships were introduced into cytoscape to construct the ceRNA network. The advantages of Cytoscape are as follows: (i) Cytoscape provides basic functions for building and querying informational networks, allowing visual integration of several variables. (ii) Cytoscape allows for extensions through plug-ins and accelerated computational analysis [17, 18].

The construction of the key IncRNA-miRNA-mRNA sub-network

We calculated the degree of all IncRNAs in the ceRNA network and selected IncRNAs with a threshold of degree > 7 [19]. All IncRNA-related miRNAs and mRNAs were extracted from...
Potential prognostic biomarkers for gastric adenocarcinoma

the ceRNA network, and the key lncRNA-miRNA-mRNA sub-network was visualized by using Cytoscape software.

Functional enrichment analysis

In order to explore biological processes and pathways involved in the mRNAs in the sub-network, GO enrichment analysis was performed using the BINGO plug-in in Cytoscape software.

Survival analysis

To further explore the prognostic value of key genes in the sub-network, we performed a survival analysis. The survival curves of the differentially expressed mRNAs, lncRNAs, and miRNAs were plotted using “survival” package, and the differences between the two groups were tested using a log-rank test. A p value < 0.05 was considered statistically significant [20].

Validation of the key genes

Gene Expression Profiling Interactive Analysis (GEPIA; http://gepia.cancer-pku.cn) was used to validate the differential expression of the hub genes between gastric adenocarcinoma and normal samples. We also analyzed the differential expression of the hub genes at different stages. Additionally, survival analysis was performed by GEPIA. Furthermore, the protein difference of hub gene was validated between gastric adenocarcinoma and normal samples by immunohistochemistry via the Human Protein Atlas database (HPA; https://www.proteinatlas.org/).

Results

Differential expression analysis of mRNAs, lncRNAs and miRNAs

The differentially expressed mRNAs, lncRNAs and miRNAs are shown in the volcano plots and heat maps. First, the differential expression analysis was performed between 32 normal samples and 373 tumor samples, resulting in 1635 differentially expressed mRNAs including 853 up-regulated and 782 down-regulated mRNAs (Figure 1A and 1D). Furthermore, 1,002 differentially expressed lncRNAs between 32 normal samples and 373 tumor samples were identified, including 791 up-regulated and 211 down-regulated lncRNAs (Figure 1B and 1E). Moreover, differentially expressed miRNAs were analyzed between 41 normal samples and 434 tumor samples. There were 98 differentially expressed miRNAs including 81 up-regulated and 17 down-regulated miRNAs (Figure 1C and 1F).

Prediction of miRNA-lncRNA by miRcode and starBase

Through differential expression analysis, differentially expressed lncRNAs and miRNAs were identified. We used miRcode and starBase databases to match the lncRNA-miRNA relationship pairs. MiRcode was used to find specific miRNA target sites in lncRNAs or potential biomarker targets for specific miRNAs. As a

<p>| Table 3. Eleven mRNAs in the sub-network |</p>
<table>
<thead>
<tr>
<th>mRNAs</th>
<th>logFC</th>
<th>logCPM</th>
<th>p value</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEST</td>
<td>2.15855444</td>
<td>5.794860908</td>
<td>9.90E-21</td>
<td>2.89E-19</td>
</tr>
<tr>
<td>SERPINE1</td>
<td>2.069614298</td>
<td>6.759560596</td>
<td>1.25E-10</td>
<td>9.25E-10</td>
</tr>
<tr>
<td>LEFTY1</td>
<td>5.141172091</td>
<td>3.567697239</td>
<td>4.58E-10</td>
<td>3.15E-09</td>
</tr>
<tr>
<td>ATAD2</td>
<td>2.01028359</td>
<td>6.000161444</td>
<td>1.21E-27</td>
<td>7.46E-26</td>
</tr>
<tr>
<td>HOXC8</td>
<td>3.865773349</td>
<td>1.457451023</td>
<td>3.56E-19</td>
<td>8.31E-18</td>
</tr>
<tr>
<td>IL11</td>
<td>3.10741378</td>
<td>2.820762254</td>
<td>5.96E-10</td>
<td>4.05E-09</td>
</tr>
<tr>
<td>TMEM100</td>
<td>-2.986552177</td>
<td>2.36180323</td>
<td>3.81E-11</td>
<td>3.62E-09</td>
</tr>
<tr>
<td>CADM2</td>
<td>-2.888739382</td>
<td>0.83770071</td>
<td>1.34E-21</td>
<td>4.43E-20</td>
</tr>
<tr>
<td>ESRRG</td>
<td>-2.9101366376</td>
<td>2.63170976</td>
<td>2.17E-31</td>
<td>2.15E-29</td>
</tr>
<tr>
<td>NPTX1</td>
<td>-2.727294831</td>
<td>3.354362995</td>
<td>4.32E-15</td>
<td>5.92E-14</td>
</tr>
<tr>
<td>CHRDL</td>
<td>-2.300010041</td>
<td>4.126124863</td>
<td>7.05E-13</td>
<td>7.27E-12</td>
</tr>
</tbody>
</table>

<p>| Table 4. Eight miRNAs in the sub-network |</p>
<table>
<thead>
<tr>
<th>miRNAs</th>
<th>logFC</th>
<th>logCPM</th>
<th>P value</th>
<th>FDR</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa-mir-122</td>
<td>3.404846093</td>
<td>2.825250337</td>
<td>0.000296106</td>
<td>0.000651851</td>
</tr>
<tr>
<td>hsa-mir-372</td>
<td>7.818666594</td>
<td>5.781891101</td>
<td>1.92E-08</td>
<td>9.95E-08</td>
</tr>
<tr>
<td>hsa-mir-373</td>
<td>7.501018965</td>
<td>3.136327419</td>
<td>3.56E-07</td>
<td>1.37E-06</td>
</tr>
<tr>
<td>hsa-mir-184</td>
<td>3.719084687</td>
<td>2.702954276</td>
<td>1.25E-07</td>
<td>5.46E-07</td>
</tr>
<tr>
<td>hsa-mir-145</td>
<td>-2.275853373</td>
<td>12.60628818</td>
<td>3.14E-34</td>
<td>2.79E-32</td>
</tr>
<tr>
<td>hsa-mir-383</td>
<td>-2.261768912</td>
<td>0.717451055</td>
<td>1.55E-14</td>
<td>1.93E-13</td>
</tr>
<tr>
<td>hsa-mir-204</td>
<td>-2.088107471</td>
<td>3.012253786</td>
<td>2.12E-15</td>
<td>2.88E-14</td>
</tr>
<tr>
<td>hsa-mir-205</td>
<td>-2.940557077</td>
<td>7.698927347</td>
<td>5.78E-09</td>
<td>3.21E-08</td>
</tr>
</tbody>
</table>
Potential prognostic biomarkers for gastric adenocarcinoma

Figure 5. The GO relationship network by use of Cytoscape plug-in BinGO (Yellow nodes: nodes with p value < 0.05, and Benjamini associated p value < 0.01).
Potential prognostic biomarkers for gastric adenocarcinoma

Figure 6. Kaplan-Meier survival curves for SERPINE1 and ADAMTS9-AS2 correlated with overall survival. (Horizontal axis: overall survival: years, vertical axis: overall survival function).

result, we identified 149 miRNA-lncRNA relationships.

The prediction of miRNA-mRNA relationship pairs by miRTarBas, miRDB, and TargetScan databases

According to our research, we identified 179 differentially expressed miRNAs and 1635 differentially expressed mRNAs. To explore miRNA-mRNA interactions, miRTarBas, miRDB, and TargetScan databases were used to predict the miRNA-mRNA relationships. Based on the three databases, we finally obtained 15 miRNA-mRNA pairs (Table 1).

The construction of the ceRNA network

Based on lncRNA-miRNA and miRNA-mRNA relationships, we built the ceRNA network (Figure 2). There were a total of 11 mRNAs, 9 miRNAs and 62 lncRNAs in the ceRNA network. Up-regulated mRNAs included MEST, SERPIN-E1, LEPTY1, ATAD2, HOXC8, IL11; and down-regulated mRNAs included TMEM100, CADM2, ESRPG, NPTX1, and CHRDL1. The ceRNA network included five up-regulated miRNAs (has-mir-122, has-mir-372, has-mir-373, has-mir-508 and has-mir-184) and four down-regulated miRNAs (has-mir-383, has-mir-145, has-mir-204 and has-mir-205). Moreover, there were 63 lncRNAs in the ceRNA network, including 43 up-regulated and 20 down-regulated lncRNAs (Table 2).

The construction of the IncRNA-miRNA-mRNA sub-network

After constructing an mRNA-miRNA-lncRNA network, we calculated the degree of all lncRNAs in the ceRNA network (Figure 3). LncRNAs with degree > 7 were screened for the sub-network. After that, we matched the IncRNA-miRNA and miRNA-mRNA associated relationships and reconstructed the key IncRNA-miRNA-mRNA sub-network (Figure 4). In the sub-network, there were 11 mRNA (Table 3), 8 miRNA (Table 4), and two lncRNA (ADAMTS9-AS2, LINC00330).

Functional enrichment analysis

To explore the biological processes involved in the 11 mRNAs in the sub-network, the GO relationship network was constructed by use of Cytoscape plug-in BinGO (p value < 0.05, and Benjamini associated p value < 0.01) (Figure 5). The results showed that these mRNAs could be widely involved in a variety of biological processes, such as cell differentiation, multicellular proliferation and inflammatory responses.

The overall survival analysis of key genes in the sub-network for gastric adenocarcinoma

Survival analysis showed that highly expressed mRNA SERPINE1 and lncRNA ADAMTS9-AS2 correlated with poor prognosis and had low expression (Figure 6). Therefore, SERPINE1
and ADAMTS9-AS2 could be related with prognosis of gastric adenocarcinoma.

Identification and validation of key genes

The mRNA SERPINE1 and IncRNA ADAMTS9-AS2 were further validated by GEPIA database. The overall survival and disease free survival revealed that highly expressed SERPINE1 and ADAMTS9-AS2 were related with worse prognosis of gastric adenocarcinoma, which was consistent with our previous study (Figure 7). Furthermore, we validated the expression of SERPINE1 and ADAMTS9-AS2 in 211 normal samples and 408 gastric adenocarcinoma samples (Figure 8A, 8B). We analyzed the expression differences of SERPINE1 and ADAMTS9-AS2 at different stages, as shown in Figure 8C, 8D. Finally, immunohistochemistry analysis also validated the expression of SERPINE1 in gastric adenocarcinoma tissue (Figure 9).

Discussion

Gastric cancer is one of the most prevalent malignant tumors in the world. The molecular mechanism of its development and metastasis have been studied by many scholars. However,
there are still many unknowns in this area of research and the research results can't be applied to clinical treatment. Therefore, it is important to explore potential new molecular markers of gastric adenocarcinoma [21-24].

First, we used the TCGA database to select the differentially expressed IncRNAs, miRNAs and mRNAs in gastric adenocarcinoma and normal samples. Then we took advantage of miRcode, starBase, and miRTarBase databases to find

**Figure 8.** Validation of SERPINE1 and ADAMTS9-AS2 expression levels in different tissues and different stages by GEPIA. A, B. The expression differences of ADAMTS9-AS2 in different tissues and different stages. C, D. The expression differences of SERPINE1 in different tissues and different stages.
Potential prognostic biomarkers for gastric adenocarcinoma

15 miRNA-mRNA pairs and 149 lncRNA-miRNA pairs. The results were used to construct the ceRNA network [25]. After calculating the node degree, we selected 2 lncRNAs to construct a lncRNA-miRNA-mRNA sub-network according to degree > 7 [26, 27]. We found 21 nodes in the sub-network, including 11 mRNA nodes, 8 miRNA nodes, and two lncRNA nodes. Two of lncRNAs were ADAMTS9-AS2, LINC00330.

In order to understand the function of the mRNAs in the sub-network, we carried out GO pathway function enriched analysis of 11 mRNAs with lncRNA in the ceRNA network. The results showed that these mRNAs may be involved in the development of cells, multicellular proliferation, and inflammatory responses and other biological processes. These biological processes are related to the development of gastric adenocarcinoma. SERPINE1 was enriched in many processes. Thus, it can be deduced that the IncRNA associated with SERPINE1 can also participate in a wide range of biological processes related to the development of gastric adenocarcinoma including ADAMTS9-AS2, LINC00330 [28, 29]. Additionally there was ADAMTS9-AS2/LINC00330-hsa-mir-145-SERPINE1 in the sub-network. In order to identify the interactions of overall survival with specific IncRNAs, miRNAs and mRNAs in the sub-network, we made the Kaplan-Meier survival analysis [30]. We finally found that SERPINE1 and ADAMTS9-AS2 were closely related to the overall survival of gastric adenocarcinoma patients. Therefore, inhibiting the expression of SERPINE1 and ADAMTS9-AS2 may prolong the survival time of patients.

Previous studies have found that overexpressed ADAMTS9-AS2 inhibits proliferation, migration and invasion of gastric cancer cells and induces apoptosis. However, the specific mechanism of action still needs further research [31]. In our study, we predicted that overexpressed ADAMTS9-AS2 could decrease patients' survival time, and be used as a potential prognostic biomarker. It has been reported that ADAMTS9-AS2 is a novel tumor suppressor of glioma regulated by DNMT1, and LncRNA ADAMTS9-AS2 may be a potential biomarker and therapeutic target for gliomas [32]. At the same time, down-regulation of ADAMTS9-AS2 is associated with poor prognosis of colorectal cancer [33]. Interestingly, ADAMTS9-AS2 is also associated with diabetic retinopathy in Chinese people with type 2 diabetes, which fully demonstrates that ADAMTS9-AS2 can participate in multiple biological pathways and will be of great significance for an in-depth study [34]. On the contrary, there are few research reports on LINC00330.

Considering ADAMTS9-AS2 could serve as a possible "sponge" of hsa-mir-145, we further analyzed the research results of hsa-mir-145. Compared with non-cancerous gastric mucosa, the expression of hsa-mir-145 is decreased in gastric cancer tissue in accordance with our results. However, research concerning hsa-mir-372 still needs further exploration. We found that mRNA SERPINE1 was up-regulated in the sub-network and highly expressed SERPINE1 could decrease the survival time of patients. SERPINE1 has been confirmed as an important regulator of tumorigenesis and is highly expressed in many cancers including gastric cancer [35, 36]. It has been reported that SERPINE1 could possess the prognostic value of gastric adenocarcinoma in accordan-
Potential prognostic biomarkers for gastric adenocarcinoma

Conclusion

We found that SERPINE1/LINC00330-hsa-mir-145-ADAMTS9-AS2 could become a possible target pathway in gastric adenocarcinoma worthy of further analysis. Inhibiting overexpressed SERPINE1 and ADAMTS9-AS2 could increase the patients’ survival time.

Acknowledgements

This work was funded by Liaoning Natural Science Foundation (20180550781).

Disclosure of conflict of interest

None.

Abbreviations

cerNA, competitive endogenous RNA; TCGA, The Cancer Genome Atlas; lncRNA, long non-coding RNA; miRNA, microRNAs; MREs, miRNA response components; ncRNAs, non-coding RNAs.

Address correspondence to: Shenyu Wang, Department of Integrated Traditional Chinese and Western Medicine, Cancer Hospital of China Medical University, LiaoNing Cancer Hospital and Institute, No. 44 Xiaoheyan Road, Dadong District, Shenyang 110042, Liaoning Province, PR China. Tel: +86-18900917654; E-mail: 148003642@qq.com

References

Potential prognostic biomarkers for gastric adenocarcinoma


