Clinical implications and mechanism of CST1 in gastric carcinoma based on database screening

Peng Liu, Jiusong Luan, Xinyu Peng, Xiu Zhang, Jianjun Zhang, Haizhe Chang

1 Department Emergency Medicine, Affiliated Hospital of Hebei University, Baoding, Hebei, 071000, China
2 Department of Respiratory and Critical Care Medicine, Affiliated Hospital of Hebei University, Baoding, Hebei, 071000, China
3 Department of Gastrointestinal Surgery, Affiliated Hospital of Hebei University, Baoding, Hebei, 071000, China

Abstract

Gastric cancer (GC) remains one of the most common malignant tumours worldwide, with extremely high morbidity and mortality rates. An in-depth understanding of the pathogenesis of GC is key to the future diagnosis and treatment of GC. In this study, we analysed the differentially expressed genes (DEGs) in gastric carcinoma (GC) through GEO database and their clinical implications, with the aim of providing clinical reference and guidance. We selected the GSE118916 dataset for bioinformatics analysis and identified a total of 3231 DEGs. Keywords, including extracellular region, vesicle, protein digestion and absorption, ECM-receptor interaction, etc., of DEGs can be seen by the GO and KEGG enrichment analysis. The online database determined up-regulated CST1 in GC and some other tumors, as well as a close connection between CST1 with patient prognosis. Subsequently, we collected a number of GC clinical cases and examined the expression of CST1, which was seen to be highly expressed in GC, with a favorable diagnostic effect on the occurrence of GC (P<0.05) and a strong correlation with TNM stage, tumor invasion, tumor diameter and differentiation (P<0.05). In other words, CST1 is closely related to the occurrence and development of GC, and has the potential to be a breakthrough in the diagnosis and treatment of GC in the future.

Keywords: Bioinformatics analysis; Gastric carcinoma; CST1; Clinical implications; Enrichment analysis

1. Introduction

Gastric carcinoma (GC) is a malignancy of gastric mucosal epithelium origin with a predilection for any part of the stomach and a prevalence second only to lung cancer among all malignant tumors[1]. Rather than any special clinical symptoms, early GC often presents similar symptoms to chronic gastric diseases such as gastritis and gastric ulcers that are easily overlooked, resulting in unsatisfactory early diagnosis of GC [2]. In China alone, GC is estimated to inflict about 60 among every 100,000, with mortality reaching up to 21.4/100,000[3]. Clinically, an in-depth understanding of the nosogenesis of GC has always been the key to finding new diagnoses and treatment schemes for the disease[4]. As relevant research advances, the onset and development of GC has been identified to be a multi-factor, multi-stage and multi-step process, involving massive molecular participation and complex network regulation[5]. As the Human Genome Project (HGP) on a global scale develops, deciphering the genetic codes of human beings and various model organisms has become an important subject in the field of biology and a hotspot in modern cancer research[6]. With the development of sequencing technology and the arrival of post-genome era, researchers believe that the identification of molecular substances such as tumor DNA (genome), mRNA (transcriptome) and protein (proteome) sequences can provide novel approaches for tumor diagnosis, clinical outcome prediction and target therapy[7]. For example, tumor cells have various DNA mutations, and finding out which genes play a key part in the early stage of tumor formation and development through bioinformatics analysis may provide data support for individualized treatment of tumors[8].

In the research of GC, the Wnt axis is the current research hotspot. This axis is a signal transduction pathway of a group of multi-downstream signal transduction pathways stimulated by the binding of the ligand-protein Wnt and membrane protein receptors, through which extracellular signals are transmitted to cells by the activation of intracellular segments of cell surface receptors[9]. It has been found that the Wnt axis is closely related to GC cell growth, development and proliferation[10, 11]. Therefore, the regulatory molecules related to the Wnt axis may be a breakthrough in future GC diagnosis and treatment.

Accordingly, this study analyzes the differentially
expressed genes (DEGs) in GC and paracancerous tissue specimens in the online database (GSE118916 dataset) through bioinformatics technology, in view of providing a new research direction for the future diagnosis and management of GC and laying a reliable foundation for the follow-up research.

2. Materials and methods

2.1. Selection of dataset
On the website of GEO (URL: https://www.ncbi.nlm.nih.gov/geo/), we entered “gastric carcinoma” and selected "homo sapiens" in the screening results to search for relevant information and then review the returning results one by one. Finally, the published dataset GSE118916, including 30 groups of samples (GC tissues and normal stomach tissues with 15 groups each), was selected for analysis on GEO2R tool (the platform is GPL15207). Results were screened based on Log (FC) >1 or < -1 and P<0.05, and volcano plot and heat map were plotted.

2.2. PPI network analysis
PPI network analysis was carried out on the screened DEGs through the STRING database (URL: https://string-db.org/) to observe potential connections between these DEGs.

2.3. Enrichment analysis
The DEGs were classified by pre-built genome annotation databases (GO and KEGG). Then, statistical algorithms were used to find out the categories (biological composition/function/process) that were significantly different from the background gene sets, and the commonality of biological composition/function/process among the specific gene sets was identified. After clustering, the redundancy was removed to obtain the gene enrichment results.

2.4. On-line expression analysis
Through GEPIA (URL: http://gepia.cancer-pku.cn/index.html) and ENCORI (URL: https://starbase.sysu.edu.cn/) databases, the expression patterns of the target genes in GC and other neoplastic diseases were analyzed to preliminarily confirm their expression patterns.

2.5. Prognostic analysis
Similarly, the prognostic influences of the target genes in GC and other neoplastic diseases were discussed through GEPIA, ENCORI and Kaplan-Meier Plotter (URL: http://kmplot.com/analysis/index.php?p=background) databases, and the survival curves drawn in the online databases were obtained.

2.6. Clinical trial
Fifty-four GC cases (research group) and sixty-two healthy people (control group) presented to our hospital between May 2021 and March 2022 were selected as the research participants for retrospective analysis. All the research subjects signed the informed consent form by themselves.

2.7. Criteria for patient enrollment and exclusion
RG: Cases enrolled presented clinical manifestations consistent with GC and were confirmed as GC[12] after biopsy in the pathology department of our hospital, with the age above 18 and complete medical records. Those with multiple tumor diseases, cardio-cerebrovascular diseases, autoimmune diseases, mental illness, liver and kidney dysfunction or a life expectancy < 1 month were excluded. CG: health controls who underwent routine physical examinations in our hospital with no previous major medical history nor abnormal physical examination results were included.

2.8. Sampling and testing
For all subjects, fasting venous blood was collected upon admission into coagulation-promoting tubes and centrifuged after standing (30 min) at room temperature to separate serum and plasma. After total plasma RNA extraction by the Trizol method and purity determination using an UV spectrophotometer, the RNA was reverse transcribed into cDNA as per kit manuals, followed by PCR reaction [13] under the conditions of 94℃ for 15s, 55℃ for 30s and 70℃ for 30s, for 40 cycles. CST1 expression, normalized with GAPDH (See Table1 for primer sequences designed and constructed by Invitrogen, USA), was calculated by 2△△CT.

2.9. Statistical processing
Statistical analysis was performed with SPSS22.0. Count data denoted by (%) was compared between groups by the Chi-square test. Measurement data were all expressed in (x±s), and the independent sample t-test was used for comparison between groups. Diagnostic value was analyzed by receiver operating characteristic (ROC) curve. The presence of significance was indicated by P<0.05.

3. Results

3.1. Analysis results of DEGs
In the GSE118916 dataset, we found a total of 3231 DEGs, among which SPP1, THBS2, SFRP4, ESRRG, CHIA, GIP, etc. were most significantly differentially expressed. The PPI network analysis revealed obvious correlation among many genes, only genes such as KIAA1522, CCDC169-SOHLH2, ZAK, and CACFD1 were less associated genes. (Fig.1).

3.2. Bioinformatics analysis results
GO analysis showed that most of the DEGs had the characteristics of extracellular region, vehicle, response to chemicals, etc. KEGG analysis also revealed the Protein digestion and absorption, ECM-receptor interaction, Complement and coagulation cascades, Phagosome, etc. in most of the DEGs. Among them, 19 genes contain the keyword "Wnt signaling pathway", namely: SFRP4/SFRP2/MMP7/PRKCB/SERPINF1/CST1/CCND2/CCDC169-SOHLH2, ZAK, and CACFD1 were less associated genes. (Fig.1).

Table1. Primer sequences.

<table>
<thead>
<tr>
<th>Gene</th>
<th>F (5’-3’)</th>
<th>R (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CST1</td>
<td>CTCGCTTCGGGACACCA</td>
<td>AACGCTTCGAATTGCGT</td>
</tr>
<tr>
<td>GAPDH</td>
<td>GAGCAGAGCCTCAGCCTTT</td>
<td>AGACAAAGACCCCGGTTT</td>
</tr>
</tbody>
</table>
RAC1/NFATC1/CAMK2G/FZD2/MYC/RUVBL1/FZD7/CACYBP/FOSL1/RSPO3/PLCB4/PRKACB. It can be seen that the above genes have potential important influences on the occurrence and development of GC (Fig.2).

3.3. Expression analysis of CST1
In the GEPIA database, CST1 was highly expressed in bladder urothelial carcinoma (BLCA), breast carcinoma (BRCA) and colon adenocarcinoma (COAD), but it presented aberrant expression in neither diffuse large B-cell lymphoma (DLBC) nor kidney renal clear cell carcinoma (KIRC). Furthermore, we investigated CST1 expression in stomach adenocarcinoma (STAD) and found obviously higher CST1 levels in 408 STAD tissue specimens compared with 211 normal tissues in the GEPIA database. Similarly, elevated CST1 expression was determined in 375 STAD samples collected in the ENCORI database compared with 32 normal samples (Fig.3).

3.4. Prognostic significance of CST1
The survival analysis based on the GEPIA database showed no significant influence of CST1 expression on the overall survival (OS) rate and disease-free survival rate of STAD patients. Nor was there any notable connection between high or low CST1 expression with patients’ OS in the ENCORI database. Among the 875 cases included in the Kaplan-Meier Plotter database, patients with high CST1 expression showed a lower OS rate than those with low CST1 expression, with the mean survival of patients with high and low expression of CST1 being 21.4 months and 36.4 months, respectively (Fig.4).

3.5. Comparison of clinical baseline data between two groups
In order to ensure the accuracy of the clinical trial, we first compared the clinical baseline data such as age, gender composition and family disease history between the research group and the control group. It can be seen that there were no statistical differences in all baseline data (P>0.05), indicating comparability between the two groups (Table 2).

3.6. Comparison of CST1 expression level
PCR results showed that the peripheral blood CST1...
mRNA of the research group was (2.32±0.58), while that of the control group was (1.78±0.67), indicating elevated CST1 expression in GC cases (P<0.001). ROC analysis exhibited that when peripheral blood CST1>1.975, the sensitivity, specificity and area under curve (AUC) for diagnosing GC were 75.93%, 66.13%, and 0.731, respectively (P<0.001) (Fig 5).

3.7. Correlation of CST1 with pathological features of GC

After analysis, it can be seen that there were no differences in CST1 expression among patients of different ages, genders, and smoking and drinking habits (P>0.05), indicating that CST1 has no significant relationship with the above indexes. However, in patients with TNM stage III-IV, tumor invasion degree T3-T4, tumor diameter >5cm and low differentiation, CST1 showed markedly increased levels (P<0.05), suggesting a close connection between CST1 and these indexes (Table 3).

4. Discussion

GC, as one of the most high-incidence malignancies in the world, has caused a huge medical burden to the clinic[13, 14]. Meanwhile, limited by the existing diagnosis and treatment conditions, the prognosis of advanced GC

**Table 2. Comparison of clinical baseline data.**

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>Gender male/female</th>
<th>Family history of disease have/none</th>
<th>Smoking yes/no</th>
<th>Drinking yes/no</th>
<th>Place of residence urban/rural</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control group</td>
<td>n=62</td>
<td>64.39±5.56</td>
<td>38/24</td>
<td>458</td>
<td>25/37</td>
<td>20/42</td>
</tr>
<tr>
<td>Research Group</td>
<td>n=54</td>
<td>65.46±5.13</td>
<td>32/22</td>
<td>549</td>
<td>24/30</td>
<td>20/34</td>
</tr>
</tbody>
</table>

**Table 3. Correlation of CST1 with pathological features of GC.**

<table>
<thead>
<tr>
<th></th>
<th>n</th>
<th>CST1 mRNA</th>
<th>t</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≤64</td>
<td>19</td>
<td>2.33±0.66</td>
<td>0.120</td>
<td>0.905</td>
</tr>
<tr>
<td>&gt;64</td>
<td>35</td>
<td>3.31±0.54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td>0.069</td>
<td>0.945</td>
</tr>
<tr>
<td>male</td>
<td>32</td>
<td>2.32±0.59</td>
<td></td>
<td></td>
</tr>
<tr>
<td>female</td>
<td>33</td>
<td>2.31±0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Family history of disease</td>
<td></td>
<td></td>
<td>0.145</td>
<td>0.885</td>
</tr>
<tr>
<td>have</td>
<td>5</td>
<td>2.35±0.41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>none</td>
<td>49</td>
<td>2.31±0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Smoking</td>
<td></td>
<td></td>
<td>0.815</td>
<td>0.419</td>
</tr>
<tr>
<td>yes</td>
<td>24</td>
<td>2.34±0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>30</td>
<td>2.30±0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Drinking</td>
<td></td>
<td></td>
<td>0.242</td>
<td>0.809</td>
</tr>
<tr>
<td>yes</td>
<td>20</td>
<td>2.29±0.52</td>
<td></td>
<td></td>
</tr>
<tr>
<td>no</td>
<td>34</td>
<td>2.33±0.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Place of residence</td>
<td></td>
<td></td>
<td>0.364</td>
<td>0.717</td>
</tr>
<tr>
<td>urban</td>
<td>42</td>
<td>2.33±0.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>rural</td>
<td>12</td>
<td>2.26±0.68</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TNM stage</td>
<td></td>
<td></td>
<td>3.621</td>
<td></td>
</tr>
<tr>
<td>I-II</td>
<td>40</td>
<td>2.16±0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>III-IV</td>
<td>14</td>
<td>2.75±0.44</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Invasion degree</td>
<td></td>
<td></td>
<td>3.225</td>
<td>0.002</td>
</tr>
<tr>
<td>T1-T2</td>
<td>42</td>
<td>2.19±0.55</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T3-T4</td>
<td>12</td>
<td>2.75±0.45</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degree of differentiation</td>
<td></td>
<td></td>
<td>2.751</td>
<td>0.008</td>
</tr>
<tr>
<td>moderately/highly differentiated</td>
<td>43</td>
<td>2.21±0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>low differentiation</td>
<td>11</td>
<td>2.72±0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tumor diameter</td>
<td></td>
<td></td>
<td>2.263</td>
<td>0.028</td>
</tr>
<tr>
<td>≤5cm</td>
<td>33</td>
<td>2.18±0.58</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&gt;5cm</td>
<td>21</td>
<td>2.53±0.51</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
patients is generally poor, with a 5-year survival rate of only 20-40\%[15]. For GC, there is increasing evidence that small molecule genes play an extremely important role. For example, miR-216b prevents GC from proliferation and migration by targeting PARK7[16], and the impaired autophagy degradation of lncRNA ARHGAP5-AS1 enhances the chemo resistance of GC[17]. By screening DEGs in GC, this study can help further understand the molecular pathogenic mechanism of GC and provide a new direction for subsequent research.

In this study, we analyzed the DEGs in GC through the GSE118916 dataset. This dataset was published on June 10th, 2019, and was completed based on the platform GPL15207 [PrimeView] Affymetrix Human Gene Expression Array, including GC tissue specimens and normal stomach tissue samples of 15 groups each, with high timeliness and accuracy. We found 200 DEGs in the GSE118916 dataset through GEO2R analysis, while GO and KEGG enrichment analyses revealed that these DEGs were mainly associated with cell proliferation and autophagy. Among them, CST1 caught our attention. CST1 is localized on human chromosome 10q11.23 and was considered as a routine substance to maintain the function of monocytes and lymphocytes at the earliest[18]. With the deepening of research, the role of CST1 in neoplastic diseases has gradually attracted clinical attention. For example, miR-375 inhibits the progression of laryngeal squamous cell carcinoma via targeting CST1[19], and silencing CST1 abrogates cancer progression and stem cell properties in papillary thyroid carcinoma[20][20]. Furthermore, CST1 is considered as a new marker of pancreatic cancer in the future[21] with great potential research value. In GC, however, despite the finding of obvious abnormal expression of CST1[22, 23], there is no further in-depth research. We found that CST1 is one of the downstream products of the Wnt axis, and its potential role is worthy of further study. As we all know, the Wnt signaling pathway, as a complex protein interaction network, is primarily implicated in embryonic development and cancer, but it also participates in normal physiological processes in adult animals[24]. At present, studies have found that in ovarian endometrioid adenocarcinoma, CST1 participates in downstream signaling of Wnt pathway and further regulates tumor cell proliferation[25]. To further confirm the relationship between CST1 and GC, we first analyzed CST1 expression through online databases. The results showed up-regulated CST1 in most tumors and GC, which was consistent with our experimental results and previous studies[26-31]. In the prognosis analysis, a connection between CST1 expression and the prognosis and survival of GC patients was identified in the Kaplan-Meier Plotter database but not in GEPIA and ENCORI databases, which may be caused by the inconsistency of information recorded in the three databases. Thus, subsequent clinical follow-up should be conducted to confirm the effect of CST1 on the prognosis of GC patients.

Finally, since no studies have confirmed the actual clinical expression of CST1 in GC, we included clinical cases for confirmation. The detection results showed markedly higher peripheral blood CST1 levels in GC patients than in controls, which is in line with above experimental results. Furthermore, the ROC analysis revealed that for GC diagnosis, CST1 had a sensitivity and a specificity of 80% and 80%, respectively, demonstrating ideal diagnostic efficiency, which also indicates that CST1 has the potential to become an auxiliary diagnostic indicator for GC in the future. Moreover, CST1 was found to be closely linked to TNM stage, tumor invasion, tumor diameter and differentiation, which shows that highly expressed CST1 mainly plays the role of oncogene in GC and can accelerate tumor progression. This also suggests that targeted silencing of CST1 may become a new option for GC treatment in the future.

However, limited by experimental conditions, we have not carried out in vitro experiments to confirm the exact influencing mechanism of CST1 on GC cells. As mentioned above, the prognostic relationship between CST1 and GC still needs to be confirmed by prognostic follow-up. In future research, we will conduct a more in-depth and comprehensive analysis of the role of CST1 in GC, so as to provide a more reliable reference for clinical practice.

5. Conclusion

CST1 shows abnormally high expression in GC and is closely related to the pathological development of GC, which may become a breakthrough in future GC diagnosis and treatment, with extremely high potential research value.

Conflict of Interests
The author has no conflicts with any step of the article preparation.

Consent for publications
The author read and approved the final manuscript for publication.

Ethics approval and consent to participate
No human or animals were used in the present research.

Availability of data and material
The data that support the findings of this study are available from the corresponding author upon reasonable request.

Funding
None.

References
6. Petryszyn P, Chapelle N, Matysiak-Budnik T (2020) Gastric Can-
Clinical significance of CST1 in gastric cancer


