75714_Auto_Edited.docx

Role of transcribed ultraconserved regions in gastric cancer and therapeutic perspectives

Gao SS *et al*. Transcribed ultraconserved regions in gastric cancer

Shen-Shuo Gao, Zhi-Kai Zhang, Xu-Bin Wang, Yan Ma, Guo-Qing Yin, Xiao-Bo Guo

Abstract

Gastric cancer (GC) is the fourth leading cause of cancer-related death. The occurrence and development of GC is a complex process involving multiple biological mechanisms. Although traditional regulation modulates molecular functions related to the occurrence and development of GC, the comprehensive mechanisms remain unclear. Ultraconserved region (UCR) refers to a genome sequence that is completely conserved in the homologous regions of the human, rat and mouse genomes, with 100% identity, without any insertions or deletions, and often located in fragile sites and tumour-related genes. The transcribed UCR (T-UCR) is transcribed from the UCR and is a new type of long noncoding RNA. Recent studies have found that the expression level of T-UCRs changes during the occurrence and development of GC, revealing a new mechanism underlying GC. Therefore, this article aims to review the relevant research on T-UCRs in GC, as well as the function of T-UCRs and their regulatory role in the occurrence and development of GC, to provide new strategies for GC diagnosis and treatment.

Core Tip: Transcribed ultraconserved region (T-UCR) is abnormally expressed in gastric cancer (GC) cells and tumors. It has been found that a variety of T-UCR affects downstream genes and related pathways, and plays a regulatory role in the proliferation, migration and invasion of GC. However, there are few relevant reviews, and this paper aims to review the related studies of T-UCR in GC. And the function of

T-UCR and its regulatory role in the occurrence and development of GC, thus providing a new strategy for the diagnosis and treatment of GC.

INTRODUCTION

Gastric cancer (GC) is one of the most common malignant tumours in the world. Although the incidence of GC has decreased, GC ranks fourth in cancer-related deaths due to its high lethality[1,2]. Due to the limited specific biomarkers of GC, patients usually have advanced GC when they are diagnosed. The tumourigenesis and development of GC is a multistep process involving countless signal transduction pathways and gene regulation. Among them, oncogenic and tumour suppressor factors, such as the transcribed ultraconserved region (T-UCR), play a key role[3,4].

In the human genome, 93% of genes can be transcribed to produce RNA, but only 2% are translated to protein. This type of RNA that lacks the potential for translation into protein is called noncoding RNA[5-9]. According to the number of bases it contains, noncoding RNA is divided into long noncoding RNAs (lncRNAs, > 200 bp) and small noncoding RNAs (< 200 bp, including rRNA, miRNA, snRNA, snoRNA, siRNA and pi-RNA)[10-13]. T-UCRs are lncRNAs, transcribed from UCRs. Recent studies have found that the expression level of T-UCRs is altered and abnormally expressed in human GC[14-17], which reveals a new mechanism for the occurrence and development of GC. Therefore, this article aims to review the relevant research on T-UCRs in GC, to further understand the specific mechanism of T-UCRs in GC cells, which will facilitate preventive measures against GC, early diagnoses, and new treatments.

ULTRACONSERVED REGION AND T-UCR

In 2004, Bejerano et al[18] compared the genomes of humans, mice, and rats. They found highly conserved DNA sequences, which were subsequently renamed ultraconserved regions (UCRs)[19]. These regions were absolutely conserved in the three species (100% identical, no insertions or deletions) and were often located in fragile sites and genomic regions of tumour-related genes. UCRs are highly conserved in the evolutionary

process, due to the long-distance enhancers and ultraconserved exons that originated from the short interspersed repetitive element retroposon family 400 million years ago. At present, such extreme conservation is still active in the Indonesian "coelacanth"[20].

UCRs represent a small part of the human genome, forming a subset of conserved sequences in intragenic and intergenic regions. UCRs are functional but do not encode proteins[21]. The length of UCRs is between 200 and 799 bp. To date, humans have found at least 481 UCR regions[18,22]. It is known that a large part of these regions can actively transcribe RNA, some of them overlap with known protein-coding sequences, and more than half are predicted to not contain any protein-coding sequences[23]. Based on their overlap with known protein-coding genes, these 481 ultraconserved elements were initially divided into three categories: Nonexons, exons and possibly exons[18,19]. However, based on the positional relationship between UCRs and genes, Mestdagh *et al*[24] reclassified them into 5 categories: Exon containing (4.2%), exonic (5%), partly exonic (5%), intergenic (38.7%) and intronic (42.6%) (Figure 1). The sense/antisense strand of each region produces two transcripts, resulting in a total of 962 possible transcripts: One corresponding to the sense genome sequence (named "+") and the other corresponding to the antisense sequence (named "+ A")[25].

T-UCRs are the transcripts of UCRs[26]. Due to their high degree of conservation, T-UCRs may have fundamental functional importance for the ontogeny and phylogeny of mammals and other vertebrates. Recent studies have shown that a T-UCR acts as a regulator in a variety of pathways (such as pri-miRNA processing, transcription regulation, translation and chromatin modification)[27-29]. It is speculated that T-UCRs may be candidate genes for cancer susceptibility because the transcription level of some UCRs is dysregulated in cancer[30,31].

T-UCRS AND CANCER

Although transcribed T-UCRs are a type of lncRNA, transcribed by UCRs, they are completely different from lncRNA. T-UCRs are absolutely conserved, while lncRNAs are the least conserved among noncoding RNA[32]. The amazing evolutionary retention

of T-UCRs strongly suggests their profound biological role in various physiological responses. As a type of evolutionarily conserved ncRNA, T-UCRs are regarded as essential for life by acting as an antisense transcription inhibitors for nearby protein coding RNA and other ncRNA. T-UCRs are thought to be involved in RNA processing or transcriptional regulation[18]. It has previously been demonstrated that T-UCRs act as regulators of gene expression[20,33]. In recent decades, increasing evidence has shown that T-UCRs are involved in carcinogenesis[23,30,31,34]. Recent studies have identified changes in T-UCR expression patterns associated with specific tumour phenotypes, such as hepatocellular carcinoma[35-37], pancreatic cancer[38,39], bladder cancer[40,41], colorectal cancer[22,26,42,43], prostate cancer[16,25,44,45], cervical cancer[46], neuroblastoma[24], breast cancer[47], lung cancer[22,48-50] and leukaemia[51], indicating a mechanism by which T-UCRs are involved in cancer development. Related studies in GC have also described changes in the expression pattern of T-UCRs[3,4,14-16], indicating that T-UCRs also play a regulatory role in the occurrence and development of GC.

T-UCRS IN GC

Uc.160+ is transcribed from the UCR on the 5q14.1 chromosome band. Honma *et al*[15] used real-time fluorescent quantitative polymerase chain reaction and in situ hybridization to detect Uc.160+. Compared with nontumour tissues, Uc.160+ expression is downregulated in GC and adenoma tissues. To further understand the biological mechanism of Uc.160+ in GC, the Uc.160+ overexpression vector was used to transfect GC cell lines MKN-1 and MKN-45, and then Western blotting was performed to detect the involvement of mitogens. Compared with the control group, the expression of phosphorylated-Akt and phosphorylated-ERK in the cell line transfected with the Uc.160+ overexpression vector decreased, while the expression of PTEN increased. These results suggest that Uc.160+ exclusively repressed the phosphorylation of Akt by regulating PTEN expression. In addition, Pang *et al*[17] found through a series of phenotypic experiments that the proliferation and activity of GC cells were inhibited after Uc.160 was overexpressed, and cell apoptosis was enhanced, which further proved

the inhibitory effect of Uc.160 on GC. In conclusion, UC.160 plays an inhibitory role in the occurrence and development of GC.

Sakamoto et al[14] analysed the expression and distribution of Uc.63+ by using qRT-PCR and in situ hybridization, and they found that all GC tissues showed high Uc.63+ expression compared with normal tissues. The expression of Uc.63+ was also elevated in GC cell lines. After overexpression of Uc.63+ was induced by transfection of Uc.63+ expression vector, GC cell proliferation was significantly enhanced. After the expression of Uc.63+ was inhibited by siRNA, the proliferation of GC cells was inhibited. These results supported the possibility that Uc.63+ had a carcinogenic effect in GC. Additionally, they found that Uc.63+ had no effect on the survival rate of cancer patients, but Uc.63+ overexpression was associated with advanced cancer and GC classification. Uc.63+ was preferentially overexpressed in diffuse GC, but not in intestinal GC. In conclusion, Uc.63+ plays a key role in the classification and progression of GC.

Goto et al[16] found that UC.416+A is overexpressed in GC compared with normal tissues by in situ hybridization. The growth of GC cells was significantly inhibited after Uc.416+A was downregulated by siRNA. These results suggest that UC.416+A, as an oncogene, plays an important role in promoting the proliferation of GC cells. Global genetic analysis using Affymetrix GeneChips showed that the most upregulated gene was IGFBP1 (insulin-like growth factor binding protein 1) when Uc.416+A was overexpressed. Compared with the corresponding nontumour gastric mucosa, IGFBP6 was significantly downregulated in GC tissue, indicating that Uc.416+A might promote the proliferation of GC cells by inhibiting IGFBP6.

In addition, the expression levels of Uc.118, Uc.158, Uc.241 and Uc.346 were found to be significantly downregulated in GC; Uc.244, Uc.249, Uc.252, Uc.261, Uc.282, Uc.283 and Uc.359 Levels were undetermined (Table 1).

REGULATORY MECHANISM OF T-UCRS IN GC

The regulatory mechanism of T-UCRs in tumourigenesis and the development of cancer is largely unclear, but current studies have found that T-UCR are involved in three different regulatory mechanisms in tumours: CpG island methylation[52], interaction with miRNAs[53-55] and direct binding to the target mRNAs[27,43]. In GC related research, it was only found that T-UCRs participate in the occurrence and development of GC through the first two pathways (Figure 2).

CpG island methylation

The field of epigenetics describes the transmission of information through heritable changes in phenotypes that do not involve changes in DNA sequences during cell division[56]. CpG island methylation, histone modification and chromatin delivery structure are the potential mechanisms of epigenetic transmission, and CpG island methylation is a key component of the changes in gene expression associated with human cancer, especially the expression of GC related genes[57-60]. GC is the most susceptible of all cancers to epigenetic changes without any changes in DNA[59,61-63]. CpG islands are DNA fragments of at least 0.5 kb, that are rich in G:C and CpG content, and are present in approximately 70% of human gene promoters. In other words, CpG islands are tandem repeats of cytosine (C) and guanidine (G), where p is the phosphoric acid between C and G. In brief, the methylation of CpG occurs under the action of methylase whereby the hydrogen on the cytosine (C) 5 carbon atom is replaced by methyl (CH3)[60,64-66]. Demethylation is the opposite process. It has been found that more than 50% of human genes are regulated by promoters including CpG islands[67]. In normal cells, except for genes with inactive X chromosomes or genes related to imprinted genes, promoter CpG islands are usually unmethylated[68]. Although the aetiology is still unclear, promoter CpG island methylation may be related to cancer development and ageing[69]. It is a common feature of human cancers that low expression of tumour-suppressive ncRNA causes CpG island methylation and then affects epigenetic silencing. Promoter CpG island methylation has been found in almost all human cancer tissue types, and it is an important mechanism for the inactivation of

tumour suppressor genes and tumour-related genes[70]. GC is the most common human cancer caused by methylation of the promoter CpG island. Interestingly, T-UCRs undergo DNA methylation-related silencing in cancer cells, and promoter CpG island methylation is now considered to be important for inactivating tumour suppressor genes or tumour-related genes. Previous studies have shown that T-UCR regulation in the occurrence and development of GC is closely related to the methylation of CpG islands of the host gene promoter[16].

Goto *et al*[16] discovered that there are CpG islands approximately 500bp upstream of the Uc.158+A transcription gene. Bisulfite genomic sequencing of GC cell lines and GC tissue samples showed specific DNA methylation of GC, which contained UCRs. The luciferase vector of the CpG island upstream of the Uc.158+A transcription gene proved that when the upstream sequence of the Uc.158+A transcription gene was methylated, reporter activity was significantly inhibited. Thus, this suggests that Uc.158+A expression is silenced by DNA methylation in the promoter region upstream of its transcription gene.

Interaction with miRNAs

Previous studies have shown that miRNAs affect the regulation of gene expression at both the transcriptional and posttranscriptional levels[71-73]. MiRNAs are also almost completely conserved[74,75]. Do T-UCRs interact with miRNAs to regulate the occurrence and development of cancer? In 2007, Calin and his collaborators demonstrated for the first time that certain T-UCRs are altered due to the direct regulation of high-level miRNAs in chronic lymphocytic leukaemia, and the expression of each T-UCR is negatively correlated with the corresponding microRNA level[19]. Recent studies have found that a T-UCR interacts with miR-596 and synergistically promotes the development of bladder cancer[28]. Terreri *et al*[55] found that the formation of T-UCR::miRNA pairs may have different effects, either targeting T-UCR or forming sponges that capture miRNAs. They also reported that the interaction between miRNAs and T-UCRs can act as a network to regulate the availability of certain lncRNAs in

bladder urothelial carcinoma cells. These results indicated that miRNAs are involved in the process of T-UCR-mediated tumour regulation. In GC, we found that T-UCRs represent the possible targets of miRNAs[14-16], and these interactions may have biological and prognostic significance for cancer patients.

Goto et al[16] determined that the expression of miR-153 in GC cell lines with higher expression of Uc.416+A was significantly reduced, suggesting an inverse correlation between Uc.416+A and miR-153. The direct interaction between Uc.416+A and miR-153 was confirmed by a luciferase activity assay. This result showed that the overexpression of Uc.416+A is related to the downregulation of miR-153 in GC. Overexpression of UC.416+A caused changes in downstream related genes. The expression of insulin-like growth factor binding protein 6 (IGFBP6) was upregulated, and alcohol dehydrogenase 1C (ADH1C), homeobox B5 (HOXB5) and homeobox B6 (HOXB6) levels were downregulated. The expression of IGFBP6, HOXB5 and HOXB6 in GC and adjacent tissues showed statistically significant differences, indicating that the overexpression of UC.416+A is regulated by miR-153, which then affects the gene changes of downstream coding proteins and plays a promoting role in the growth of GC.

CpG island methylation and miRNAs play a role together

Honma et al[15] found that the low expression of Uc.160+ in GC was caused by methylation of its upstream promoter region, and the low expression of Uc.160+ caused a decrease in PTEN expression, an increase in phosphatase activity and activation of the MAPK pathway. In addition, Pang et al[17]'s study found that UC.160 was also regulated by miR-155, which then regulated PTEN and affected the MAPK pathway. In other words, the low expression of UC.160 in GC is regulated by miR-155 and affected by methylation in the upstream promoter region of its transcriptional genes.

Other factors affecting T-UCR transcription

Sakamoto et al[14] found that the expression of UC.63+ is regulated by hypoxia. The promoter region located upstream of the Uc.63+ transcription gene is considered to

have hypoxia-induced binding sites. Hypoxia induces overexpression of UC.63+, and the overexpression of UC.63+ upregulates its downstream target P65 and activates the NF-KB signalling pathway to promote the occurrence of GC (Figure 3).

CONCLUSION

In GC, T-UCRs change at the transcriptional level, and the abnormal expression of T-UCRs can lead to the occurrence and development of GC. Compared with normal cells, T-UCRs in GC cells have a unique expression profile, which indicates that changes in T-UCRs are involved in this malignant process. This study may provide new ideas and directions for GC diagnosis and prognosis. Compared with coding RNA, T-UCRs have incomparable advantages. Because T-UCRs do not code for proteins, they are relatively less regulated and more accurate. In addition, some T-UCR expression levels are completely different in different stages of cancer, and can also be used as one of the criteria for judging prognosis.

Despite the fact that this new type of dysregulated molecule seems to be useful in future clinical applications, further research is needed before it can be used as a valuable clinical biomarker for GC. First, although approximately 98% of the entire genome contains nonprotein coding genes, the proportion of T-UCRs is currently unclear. Second, current screening methods for T-UCRs (RT-qPCR, T-UCR chip) are useful, but simpler and more practical techniques will help in the identification and screening of disease-related T-UCRs. Third, it has been proven that T-UCRs are involved in the occurrence and development of GC, but their biological mechanism has not been fully elucidated. In addition, more powerful strategies are needed to clarify the regulatory role of T-UCR by constructing an interaction network, or to evaluate its function in typical signalling pathways. Since the knowledge of T-UCR is still in its infancy-especially when compared to other ncRNAs, further research is needed to convincingly incorporate these ncRNAs into the growing field of cancer therapeutics.

At present, there are few studies on T-UCR therapy, so there is still much room to explore treatment measures aimed at T-UCRs. Perhaps we can achieve a treatment

purpose by changing the expression of T-UCRs. In the reference cell experiment, restoring the downregulated levels of T-UCRs or suppressing the overexpressed levels of T-UCRs *via* overexpression vectors or small interfering RNA methods can reverse the tumour phenotype. In clinical treatment, we can use the negative regulation between T-UCRs and miRNAs to inhibit the expression of T-UCRs, and methylation inhibitors can be used to restore the expression of T-UCRs with downregulated levels in tumours, thereby delaying or even reversing tumour progression.

Although these hypotheses have not been confirmed, we believe that as research on T-UCRs continues, the insight gained will definitely provide a novel strategy for the diagnosis, prognosis and treatment of tumours.

75714_Auto_Edited.docx

Internet

19 words — 1%

8 journals.lww.com
   Internet

17 words — 1%

9 Goto, K, S Ishikawa, R Honma, K Tanimoto, N
   Sakamoto, K Sentani, N Oue, J Teishima, A
   Matsubara, and W Yasui. "The transcribed-ultraconserved
   regions in prostate and gastric cancer: DNA hypermethylation
   and microRNA-associated regulation", Oncogene, 2015.
   Crossref

13 words — < 1%

10 George Haddad, Malte Kölling, Urs A. Wegmann,
   Angela Dettling et al. " Renal AAV2-Mediated
   Overexpression of Long Non-Coding RNA Attenuates Ischemic
   Acute Kidney Injury Through Sponging of microRNA-30a-5p ",
   Journal of the American Society of Nephrology, 2021
   Crossref

12 words — < 1%

11 Yuzuru Satake, Yuki Kuwano, Tatsuya Nishikawa,
   Kinuyo Fujita et al. "Nucleolin facilitates nuclear
   retention of an ultraconserved region containing
   <i>TRA2β4</i> and accelerates colon cancer cell
   growth", Oncotarget, 2018
   Crossref

11 words — < 1%

12 Jiang Chen Peng, Jun Shen, Zhi Hua Ran.
   "Transcribed ultraconserved region in human
   cancers", RNA Biology, 2014
   Crossref

10 words — < 1%

13 www.frontiersin.org
   Internet

10 words — < 1%