Long non-coding RNAs in gastric cancer: versatile mechanisms and potential for clinical translation

Jing Zhao1*, Yongchao Liu2*, Guangjian Huang2, Peng Cui1, Wenhong Zhang1, Ying Zhang1,3

1Department of Infectious Diseases, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai 200040, China; 2Department of General Surgery, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai 200040, China; 3Department of Molecular Microbiology and Immunology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, Maryland 21205, USA. *Equal contributors.

Received January 6, 2015; Accepted February 1, 2015; Epub February 15, 2015; Published March 1, 2015

Abstract: Gastric cancer (GC) remains a serious threat to many people, representing the second leading cause of cancer-related death worldwide. The lack of early diagnostic biomarkers, effective prognostic indicators and therapeutic targets all account for the poor prognosis of GC. Therefore, the identification of novel molecular biomarkers for early diagnosis, therapeutic response, and prognosis are urgently needed. High-throughput sequencing has identified a large number of transcribed long non-coding RNAs (IncRNAs) throughout the human genome. Accumulating evidence demonstrates that these IncRNAs play multiple roles in regulating gene expression at the transcriptional, post-transcriptional, and epigenetic levels. Aberrant expression of IncRNAs occurs in various pathological processes, including GC. Many dysregulated IncRNAs in GC have been significantly associated with a larger tumor size, higher degree of tumor invasion, lymph node and distant metastasis, and poorer survival outcome. In this review, we will provide an overview of the pathogenesis of GC, the characteristics and regulatory functions of IncRNAs, and the versatile mechanisms of IncRNAs in GC development, as well as evaluate the translational potential of IncRNAs as novel diagnostic and prognostic biomarkers and therapeutic targets in GC.

Keywords: Gastric cancer, IncRNA, diagnosis, prognosis, therapy, biomarker

Introduction

Gastric cancer (GC) is one of the most common malignancies and represents the second leading cause of cancer-associated mortality worldwide [1, 2]. Many factors are implicated in the carcinogenesis of GC, including genetic factors, Helicobacter pylori infection, unhealthy diet (for example, high intake of salt and nitrates), and smoking [3-6]. Most GC cases are diagnosed at an advanced stage due to a lack of typical early symptoms. Presently, surgical resection and chemoradiotherapy are the main treatment approaches for GC [7], but relapse, distant metastasis and chemo-resistance frequently occur, and the overall 5-year survival for GC is only approximately 25% [8]. Despite great efforts to understand the biological properties of cancer, there have been minimal improvements in the clinical outcome of GC. The lack of diagnostic biomarkers, prognostic indicators and effective therapeutic targets account for the poor outcome of GC. Thus, the underlying mechanisms of GC must be interrogated so that novel promising diagnostic and prognostic molecular biomarkers can be identified and developed to improve the quality of life and survival of GC patients.

Comprehensive investigations over the past few decades have mainly concentrated on the function of critical protein-coding genes and genome alterations in the pathogenesis of GC [9, 10]. Recent progress in sequencing technology and genome-wide analysis has revealed that protein-coding genes only account for less than 2% of the total transcriptome, whereas most genes are transcribed to non-protein-coding RNAs (ncRNAs) [11]. Small ncRNAs have lengths shorter than 200 nucleotides (nt) and include microRNAs (miRNAs), small-interfering RNAs (siRNAs) and Piwi-interacting RNAs (piRNAs), and their functions and regulatory mechanisms have been extensively studied, especially
Long non-coding RNAs in gastric cancer

the miRNAs [12]. Through complete or incomplete base-pairing complementary to the 3′-untranslated regions (3′-UTRs) of mRNAs, miRNAs mediate targeted mRNA degradation or translational inhibition [13, 14]. As the regulators of gene expression, miRNAs play important roles in cancer, and various dysregulated miRNAs are associated with GC [15, 16]. In addition to these small ncRNAs, long ncRNAs (lncRNAs) with lengths of more than 200 nt are abundant in the human genome and have attracted increasing scientific interest. Despite being initially regarded as “transcriptional noise”, accumulating evidence has found that lncRNAs can manipulate local or global gene expression via transcriptional, post-transcriptional and epigenetic regulation [17]. The lncRNAs that have been characterized are implicated in diverse physiological and pathological processes, such as X-chromosome inactivation, stem cell pluripotency, development, immune response, cell differentiation, apoptosis, and cancer metastasis and invasion. lncRNAs can interact with proteins, DNA and RNA transcripts to control alternative splicing, chromosome remodeling, nuclear import and mRNA decay, and IncRNAs participate in almost every aspect of gene expression programs [18]. Aberrant lncRNA expression has been discovered in many types of cancers, including GC [19, 20]; these dysregulated IncRNAs can function as oncogenes or tumor suppressors to alter cellular pathways [21]. Some IncRNAs are expressed in disease- or tissue-specific patterns, which make IncRNAs attractive as diagnostic, prognostic and therapeutic biomarkers in cancers.

In this review, we summarize the pathogenesis of GC and the functional activity of IncRNAs, especially the underlying mechanisms of IncRNAs in GC development. The aim is to evaluate the significant potential of IncRNAs as novel biomarkers for early diagnosis or prognosis or as therapeutic targets in GC.

Pathogenesis of GC

Environmental risk factors of GC development

GC remains an ongoing serious threat to the public health. There are an estimated 989,000 new cases and approximately 738,000 patients dying from GC throughout the world annually [2, 22]. Although GC is the fourth most common malignancy, it is the second most frequent cancer-related cause of death worldwide [2, 22]. GC carcinogenesis is a complex and multistep process with high molecular heterogeneity. Environmental factors and genetic alterations within the host play important roles in the etiology of GC development.

Stomach infections by the gram-negative, microaerophilic bacterium H. pylori is a well-established risk factor for GC [23]. The World Health Organization (WHO) defined H. pylori as a class I carcinogen in 1994, and approximately 50% of the population are infected [24]. H. pylori infection can induce chronic inflammation, reactive oxygen species (ROS) accumulation and DNA damage in the gastric mucosa, which induces the normal gastric epithelium to develop atrophic gastritis, intestinal metaplasia, and dysplasia, with eventual progression to carcinoma [25]. Furthermore, H. pylori infection has been found to enhance aberrant promoter methylation modification in the gastric mucosa and gastric epithelial cells, leading to the silencing of certain tumor suppressor genes and the promotion of gastric carcinogenesis [26-28]. Clinical studies suggest that H. pylori eradication can effectively reduce precancerous lesions and GC [29]. In addition to H. pylori infection, Epstein-Barr virus (EBV) infection has also been found to cause GC development [30].

Unhealthy dietary habits and lifestyles are also important factors that increase the risk of GC development. Epidemiological data suggest that a diet rich in salt, N-nitroso compounds and fat is a risk factor for GC development [31]. Salt and N-nitroso ingestion can cause mechanical damage to the gastric mucosa, induce gastritis and promote persistent H. pylori infection [32]. Fresh fruits and vegetables are full of carotenoids, vitamin C, folate, phytochemicals, and fiber, which can modestly reduce GC susceptibility [33]. Cigarette smoking and alcohol intake are established risk factors for GC [34]. Alcohol can stimulate the gastric mucosa, and tobacco may induce precursor gastric lesions and increase H. pylori infection. Drinking green tea may prevent carcinogenesis, as green tea is the most abundant source of epigallocatechin gallate (EGCG), which protects gastric epithelial cells from H. pylori-induced cytotoxicity [35].

Molecular mechanisms of GC based on mRNA and protein alterations

Most GC patients suffer from malignant gastric epithelial lesions. However, GC has highly het-
Long non-coding RNAs in gastric cancer

erogeneous properties that can be classified into various subgroups based on histological, anatomical, epidemiological and molecular characteristics. Currently, according to gene expression profiling with consensus hierarchical clustering selections, three major subtypes of GC have been identified: proliferative, metabolic and mesenchymal, each of which exhibits characteristic variations in molecular and genetic properties and responses to chemotherapy [36]. Proliferative GC presents high levels of genomic instability, including tumor suppressor p53 (TP53) mutations and DNA methylation. Metabolic patients appear to respond better to 5-fluorouracil treatment. The mesenchymal subtype may have cancer stem-like cells that are sensitive to inhibitors targeting the PI3K/AKT/mTOR signaling pathway [36]. These molecular classifications for GC provide the rationale for more effective personalized therapy.

With advances in high-resolution sequencing technology, a wide range of somatic alterations have been studied in GC [10]. There are frequent gene mutations in TP53, PIK3CA and ARID1A. Mutations in cell adhesion genes, such as FAT4 and the chromatin remodeling genes ARID1A, MLL3 and MLL, are a common occurrence in GC [10]. Among the numerous gene mutations implicated, TP53 alterations have been extensively studied. TP53 plays important roles in cell fate determination and has been described as “the guardian of the genome”. Inactivation mutations and loss of heterozygosity (LOH) of TP53 occur at high frequency and seem to be early events in GC [37].

A comprehensive genomic analysis of 233 GC samples revealed that receptor tyrosine kinase (RTK)/RAS alterations collectively occurred in up to 37% of GC patients, mainly comprising alterations in FGFR2, KRAS, EGFR, HER2 and MET [38]. HER2 is a well-characterized oncogene with tyrosine kinase activity that belongs to the EGFR family. The dimerization of HER2 can induce the autophosphorylation of tyrosine residues within the cytoplasmic domain, which can activate many cellular signaling cascades to trigger proliferation and carcinogenesis. Amplification or overexpression of HER2 is found in approximately 10-30% of gastric/gastroesophageal cancers and could serve as a prognostic and predictive marker. Trastuzumab monoclonal antibody directed at HER-2, which has been shown to decrease shedding of the HER-2 extracellular domain and reduce its dimerization, has now become a standard first-line treatment option for advanced HER-2-positive GC [39].

Our previous studies found several examples of aberrant expression of mRNAs and proteins in GC, which play vital roles in GC development. For example, VEGF-C and CNTN1 levels were significantly correlated with tumor size [40]; overexpressed JMJD2A was positively associated with tumor stage, nodal status and poor prognosis; and PCBP2 upregulation was linked to shorter survival time in GC [41, 42]. JMJD2A can catalyze the demethylation of histone H3 lysines 9 and 36, and PCBP2 is an RNA-binding protein. Our findings suggested that both JMJD2A and PCBP2 could reduce the expression of pro-apoptotic miR-34a to promote cell proliferation and suppress apoptosis in GC cells [41, 42]. Recently, the upregulated oncogenic protein AEG-1 was shown to associate with proinflammatory signaling in GC. Bacterial lipopolysaccharide (LPS) could induce AEG-1 expression, enhancing nuclear translocation of the NF-κB p65 subunit and, in turn, decreasing the TLR4 negative regulator SOCS. Eventually, AEG-1 promotes uncontrolled inflammation in the GC microenvironment and aggravates malignant progression [43]. AEG-1 depletion has been found to inhibit invasion and the epithelial-mesenchymal transition (EMT) program in cervical cancer [44] and to suppress cell migration in hepatocellular carcinoma [45]. In GC, AEG-1 also exerts critical roles and is a potential prognostic biomarker and therapeutic target [43]. Additionally, several single-nucleotide polymorphisms (SNP) have been associated with GC risk in genome-wide association studies (GWASs). For example, SNPs at 8q24.3 located in PSCA and SNPs at 1q22 within MUC1 are significantly associated with GC susceptibility according to three independent GWASs: two studies that examined a Chinese population [46, 47] and one that examined Japanese and Korean populations [48].

Many diagnostic, prognostic and therapeutic biomarkers have been developed. For example, serum-based carcinoembryonic antigens (CEA) and tissue-based HER2 have been suggested as potential GC markers and used in clinical practice. However, low sensitivity and specificity constrain their application. So far, there are no good markers for early GC detection and
prognosis. This situation indicates that our current knowledge of the complexities of GC remains limited, and novel pathogenic mechanisms should be explored to identify novel biomarkers for better clinical applications in GC.

**The versatile long non-coding RNAs**

Over the past decade, gene-tiling array and RNA deep sequencing (RNA-seq) studies have revealed that the human genome is pervasively transcribed and produces thousands of transcripts that lack obvious coding capacities [11]. These non-coding transcripts were initially regarded as genomic noise, but emerging evidence has demonstrated that the proverbial “dark matter” actually has important effects on the regulation of gene expression [12]. These findings have greatly challenged the conventional dogma that only proteins can perform these cellular functions and that RNA serves only as a template between DNA sequences and proteins. In the past, intensive efforts have concentrated on the functions of proteins in the pathogenesis of cancer. Recently, non-coding RNAs have attracted more attention, not only shorter transcripts (< 200 nt, which include siRNAs, piRNAs, and miRNAs and have been well documented) but also long non-coding RNAs (lncRNAs) that possess more complex regulatory mechanisms and take part in multiple biological processes, including cancer [21].

**Biological characteristics of lncRNAs**

LncRNAs are commonly defined as transcripts with lengths ranging from 200 nt to 100 kb that have little or no protein coding capacity. LncRNAs share many characteristics with mRNAs, are transcribed by one of three DNA-dependent RNA polymerases (Pol I, II, or III), have 5’ capping and polyadenylation, and undergo alternative splicing. The tertiary structures play important roles in lncRNA functions [17]. LncRNAs can be classified into diverse subgroups based on various criteria, which may help to elucidate their regulatory mechanisms. According to their functional relevance, lncRNAs can be divided into “housekeeping” and “regulatory” lncRNAs. The “housekeeping” lncRNAs are constitutively expressed, such as the tRNA and rRNA involved in protein biosynthesis. Regulatory lncRNAs may display dysregulated expression under certain physiological and pathological conditions. Based on their orientation and locations relative to neighboring genes, lncRNAs can be further classified as sense, antisense, divergent, convergent, intronic and intergenic [11]. The five broad categories of lncRNAs are described as follows: (1) “sense” and (2) “antisense” describe lncRNA transcripts that overlap with one or more exons of another transcript on the same or opposite strands, respectively; (3) “bidirectional” indicates that lncRNA expression and the neighboring coding transcript on the opposite strand is initiated in close genomic proximity; (4) “intronic” indicates a lncRNA derived from the intron of a second transcript; and (5) “intergenic” describes a lncRNA that lies within the genomic intervals between two genes as an independent unit. Additionally, some lncRNAs can modulate the expression of neighboring genes within the same chromosome in a cis manner or control distant gene expression located on the same or different chromosomes in a trans manner by affecting RNA polymerase complex recruitment or chromatin remodeling. Another classification system is based on the role of lncRNA in cancer development and comprises oncogenic lncRNAs and tumor suppressor lncRNAs.

LncRNA subcellular localization has been studied using RNA fluorescence in situ hybridization (FISH). Identifying the subcellular localization of lncRNAs provides some insight into their functional and potential interacting partners. Similar to mRNA, lncRNAs have been shown to localize across a wide range of subcellular organelles. A subset of lncRNAs are selectively localized in the nucleus [49-51], some have been visualized specifically in the cytoplasm [52], and some seem to appear in both the nuclear and the cytoplasmic compartments [53].

The evolutionary conservation of lncRNAs has been compared across species. Some lncRNAs exhibit a high level of nucleotide sequence identity within vertebrates, as exemplified by MALAT1 and its murine ortholog hepcarcin [54]. Moreover, certain lncRNAs are derived from ultra-conserved genomic regions (UCR) and are fully conserved among orthologous regions in human, rat, and mouse genomes [55]. A subset of lncRNAs show tissue-specific expression patterns in equivalent genome loci across species despite the absence of a conserved nucleotide sequence [56].
## Long non-coding RNAs in gastric cancer

<table>
<thead>
<tr>
<th>Database name</th>
<th>Description</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>LncRNA2Target</td>
<td>A database for differentially expressed genes after IncRNA knockdown or overexpression</td>
<td><a href="http://www.lncrna2target.org">http://www.lncrna2target.org</a></td>
</tr>
<tr>
<td>LNCipedia</td>
<td>A database for annotated human IncRNA transcript sequences and structures</td>
<td><a href="http://www.lncipedia.org">http://www.lncipedia.org</a></td>
</tr>
<tr>
<td>IncRNAdb</td>
<td>Expanding the reference database for functional long noncoding RNAs</td>
<td><a href="http://incrnadb.org">http://incrnadb.org</a></td>
</tr>
<tr>
<td>LncRNAWiki</td>
<td>Harnessing community knowledge in collaborative curation of human long non-coding RNAs</td>
<td><a href="http://incrna.big.ac.cn">http://incrna.big.ac.cn</a></td>
</tr>
<tr>
<td>LncRNADisease</td>
<td>A database for long-non-coding RNA-associated diseases</td>
<td><a href="http://cmbi.bjmu.edu.cn/lncrnadisease">http://cmbi.bjmu.edu.cn/lncrnadisease</a></td>
</tr>
<tr>
<td>IncRNASNP</td>
<td>A database of SNPs in IncRNAs and their potential functions in human and mouse</td>
<td><a href="http://bioinfo.life.hust.edu.cn/IncRNASNP/">http://bioinfo.life.hust.edu.cn/IncRNASNP/</a></td>
</tr>
<tr>
<td>DIANA-LncBase</td>
<td>Experimentally verified and computationally predicted microRNA targets on long non-coding RNAs</td>
<td><a href="http://www.microrna.gr/LncBase">www.microrna.gr/LncBase</a></td>
</tr>
<tr>
<td>NONCODE</td>
<td>Exploring the world of long non-coding RNA genes</td>
<td><a href="http://www.bioinfo.org/noncode/">http://www.bioinfo.org/noncode/</a></td>
</tr>
<tr>
<td>ChIPBase</td>
<td>A database for decoding the transcriptional regulation of long non-coding RNA and microRNA genes from ChIP-Seq data</td>
<td><a href="http://deepbase.sysu.edu.cn/chipbase/">http://deepbase.sysu.edu.cn/chipbase/</a></td>
</tr>
<tr>
<td>starBase v2.0</td>
<td>Decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data</td>
<td><a href="http://starbase.sysu.edu.cn/">http://starbase.sysu.edu.cn/</a></td>
</tr>
<tr>
<td>Rfam</td>
<td>The universally acclaimed database of RNA families, as well as several databases on long non-coding RNA, microRNA and their targets</td>
<td><a href="http://rfam.sanger.ac.uk">http://rfam.sanger.ac.uk</a></td>
</tr>
<tr>
<td>Linc2GO</td>
<td>Functional annotation of human lincRNA based on the ceRNA hypothesis</td>
<td><a href="http://www.bioinfo.tsinghua.edu.cn/~liuke/Linc2GO/index.html">http://www.bioinfo.tsinghua.edu.cn/~liuke/Linc2GO/index.html</a></td>
</tr>
<tr>
<td>circBase</td>
<td>Database of circular RNAs</td>
<td><a href="http://www.circbase.org">http://www.circbase.org</a></td>
</tr>
</tbody>
</table>
All of the abovementioned features imply the functional importance of lncRNAs. Several lncRNAs have been annotated; however, only a few of them have well-characterized functions. Multiple lncRNA databases have been constructed to understand lncRNAs more systematically. Here we summarized them in Table 1.

**Mechanisms of lncRNA-mediated regulation of gene expression**

LncRNAs play important roles in regulating gene expression via various mechanisms and at multiple levels. LncRNAs have the capacity to interact with proteins, DNAs or RNAs to act versatile functions. Moreover, certain lncRNAs can also function as the precursors for small RNAs to produce mature miRNAs.

**LncRNA interacts with proteins:** LncRNA can serve as a scaffold to recruit protein complexes and influence gene expression. Recent studies have revealed that a lncRNA known as HOTAIR can bind polycomb repressive complex 2 (PRC2) with its 5' domain and interact with the LSD1/CoREST/REST complex with its 3' end. PRC2 is composed of methylase EZH2, SUZ12 and EED, which are responsible for histone H3 lysine 27 trimethylation (H3K27me3); whereas the LSD1 complex serves as the demethylase that mediates histone H3 lysine 4 demethylation (H3K4me2). HOTAIR mediates two distinct complex assemblies, enabling PRC2 and LSD1 to target specific gene sites to induce H3K27me3 and H3K4me2 modification and ultimately repress gene transcription across 40 kb of the HOX D locus, despite the fact that HOTAIR itself is transcribed from the HOX C region [57, 58]. The aberrant gene expression induced by HOTAIR increases the invasiveness and metastasis of cancer [57, 59, 60]. Overexpression of HOTAIR has been observed in various cancer types, such as breast, hepatocellular, gastric, colorectal and pancreatic cancers, and affects patient survival and prognosis [59].

LncRNA can also regulate nuclear trafficking. Nuclear factor of activated T cells (NFAT), a sensitive transcriptional factor response to the alteration of calcium signals, is essential for immune response mediated by the T cell receptor [61]. When intracellular calcium levels increase, the calcium-mediated phosphatase calcineurin dephosphorylates NFAT complex subunits in the cytoplasm, which promotes NFAT translocation into the nucleus to become transcriptionally active. The lncRNA NRON was identified in a complex with nuclear import factors that specifically modulated the nuclear trafficking of NFAT. NRON inhibited the nuclear accumulation of NFAT, which prevented the active transcription of NFAT, although the precise mechanisms are unclear [62]. The lncRNA 5′aHIF-1α is distributed at the perinuclear area of kidney cancer cells and co-localizes with the nuclear pore complex protein Nup62. Antitumor inhibitor camptothecin (CPT) treatment can elevate the level of 5′aHIF-1α and is associated with a decrease in HIF-1α mRNA levels. Localization within the perinuclear area and interaction with the nuclear pore complex suggests that 5′aHIF-1α may be involved in the export of novel mRNAs into the cytoplasm [63].

LncRNAs are involved in alternative splicing via interactions with splicing factors. The lncRNA MALAT1, also known as NEAT2, is upregulated in many solid cancers and is correlated with cancer metastasis and recurrence [64]. MALAT1 is highly conserved in mammals and is localized to nuclear speckles. MALAT1 interacts with splicing factors, including the serine/arginine (SR) splicing proteins, and decreases the cellular levels of active SR splicing proteins, which repress the association of SR splicing factors with pre-mRNA [65]. A recent study found that MALAT1 could promote cell cycle progression by enhancing expression of the oncogenic transcription factor B-MYB. MALAT-1 mainly induced proliferation by attenuating the affinity of splicing factors to B-MYB pre-mRNA, leading to an aberrant alternative splicing process [66].

**LncRNA interacts with DNA, RNA or miRNA:** Natural antisense lncRNA can form an RNA duplex with its sense mRNA, which may alter mRNA stability and translation efficiency. The overlapping region of the RNA duplex may protect the mRNA from endo- or exonucleolytic degradation [67]. Moreover, antisense lncRNAs may mask the miRNA-targeting site on sense mRNA and prevent the mRNA degradation by miRNA. The BACE1 natural antisense transcript BACE1-AS increases the stability of BACE1 mRNA through the mechanisms described above [68]. This RNA duplex formation has the capacity to affect epigenetic silencing. The INK4b-ARF-INK4a locus in the human genome
Long non-coding RNAs in gastric cancer

encodes p15INK4b, p14ARF and p16INK4a, which are three tumor suppressors known to inhibit malignant cell proliferation and promote senescence and apoptosis. The natural antisense lncRNA ANRIL can form heterochromatin with INK4b-ARF-INK4a transcripts to recruit the polycomb repressor complexes PRC1 and PRC2 and establish repressive epigenetic marks on the chromatin, thereby silencing these tumor suppressors in prostate cancer [69, 70]. LncRNA can mediate ribosomal DNA (rDNA) silencing by binding and methylating the rDNA promoter region. Mammalian genomes have several clusters of tandem rDNA, most of which are silenced by heterochromatic histone modifications and CpG methylation within the rDNA promoter. RNA polymerase I produces IncRNA transcripts from rDNA promoters (termed pRNA, promoter-associated RNA). As a subset of IncRNA, pRNA is complementary to the rDNA promoter and interacts with transcription factor TTF-I to form a DNA:RNA triplex that is preferentially recognized by DNMT3b to induce CpG methylation of the rDNA promoter and silence rDNA [71].

Many studies have revealed that some IncRNAs can act as miRNA sponges by competitively interacting with miRNAs to reduce miRNA availability to their target mRNAs. The tumor suppressor gene PTEN has a pseudogene IncRNA PTENP1. PTENP1 can act as a “decoy” to sequester the miRNAs that target PTEN, thus protecting PTEN from silencing by miRNAs and exerting a growth-suppression function; however, the PTENP1 locus is selectively lost in human cancer [72].

LncRNA serves as a miRNA precursor: Several unannotated IncRNAs have the potential to produce natural precursors for miRNA-like small RNAs. For example, the imprinted IncRNA H19 has been discovered as a precursor for miR-675 [73]. miR-675 is derived from the first exon of H19 and is involved in controlling the expression of developmental genes. Additionally, H19-derived miR-675 has an impact on tumorigenesis and tumor progression by targeting tumor suppressor RB in colorectal cancer [74] and by silencing tumor suppressor RUNX1 in GC [75]. The IncRNA PVT1 locus located on 8q24 is amplified in multiple cancers. The PVT1 locus has been found to encode several miRNAs, such as miR-1204, and these miRNAs are important for T lymphomagenesis [76].

LncRNAs are associated with cancers

Cancer is a multistep process in which normal cells progressively evolve to a neoplastic state by acquiring particular capacities to disturb cellular homeostasis. Weinberg et al. proposed the following hallmarks of the malignant transformation process of cancer: (1) sustaining proliferative signaling; (2) evading growth suppressors; (3) resisting cell death; (4) enabling replicative immortality; (5) inducing angiogenesis; and (6) activating invasion and metastasis [77]. Many IncRNAs have been found to be dysregulated in cancers, and IncRNAs play important roles in each of the hallmarks of cancer [21]. Yang et al. reported that two IncRNAs, PRNCR1 and PCGEM1, which are highly overexpressed in aggressive prostate cancers could strongly enhance androgen receptor (AR)-induced transcriptional activation programs to promote cell proliferation. Their study involved an intricate series of events leading to the formation of a complex among PRNCR1, PCGEM1 and AR. PRNCR1 first bound to the C-terminally acetylated AR on the enhancers and recruited the DOT1L enzyme to methylate the N-terminus of AR, which was required for the association of PCGEM1 with AR. Subsequently, PCGEM1 interacted with the Pygo2 protein, which recognized and bound the methylated histone H3 marker H3K4me3 at the gene promoter. These successive complex associations eventually facilitated the formation of a “loop” between the enhancer and promoter sequences, which resulted in transcriptionally activated AR-targeted genes [78, 79]. The regulatory patterns of IncRNA make cancer pathogenesis more intricate and complicated, but the multi-functional and tissue-specific properties of IncRNAs provide new avenues for the development of novel diagnostic, prognostic and therapeutic biomarkers for cancers.

LncRNAs play important roles in GC

There has been increasing interest in the role of IncRNAs in GC pathogenesis. Recently, several research teams have performed IncRNA microarray profiling or RNA-seq analysis and identified many dysregulated IncRNAs in GC [20, 80-82]. These aberrantly expressed IncRNAs play critical roles in gastric carcinogenesis and aggressive progression. The upregulated IncRNAs may function as oncogenes to expedite the acquisition of the malignant cancer hallmarks, whereas the downregulated
Long non-coding RNAs in gastric cancer

IncRNAs may possess tumor suppressor features in GC.

The oncogenic functions of IncRNAs in GC

A recent study found that the IncRNA GAPLINC (gastric adenocarcinoma predictive long intergenic noncoding RNA) was highly expressed in GC specimens according to in situ hybridization (ISH) analysis [83]. The aberrant expression of GAPLINC strongly correlated with alterations in CD44, and CD44 is a well-known cancer stem cell marker that drives cancer proliferation, migration and angiogenesis. The cell migration and proliferation functions of GAPLINC could be attenuated by CD44 repression. The in vitro and in vivo data demonstrated that GAPLINC acted as a molecular decoy for miR211-3p to protect CD44 from degradation by miR211-3p [83]. In fact, there is intricate crosstalk among IncRNAs, miRNAs and mRNAs to control gene expression in GC. LncRNA can serve as a competing endogenous RNA (ceRNA) to antagonize the repressive role of miRNAs on their target mRNAs. According to IncRNA microarray data, the bioinformatics algorithm miRcode, and the miRNA targets database TarBase, the first IncRNA-miRNA-mRNA networks have been constructed for GC [84].

HOTAIR is one of the most extensively studied oncogenic IncRNAs and is upregulated in various cancers, including GC. HOTAIR overexpression has been correlated with tumor stage, venous invasion, lymph node metastasis, peritoneal dissemination, and poor overall survival rate in GC [60, 85-87]. Ectopic expression of HOTAIR promoted proliferation, the EMT program, and migration and invasion of GC cells, while HOTAIR knockdown effectively inhibited these malignant phenotypes, reduced the expression of MMP1, MMP3 and Snail proteins, and suppressed tumor growth and peritoneal metastasis in the xenograft mouse model [85, 87]. The well-studied mechanism for HOTAIR is mediated through interaction with the PRC2 and LSD1/CoREST/REST complexes, thus leading to gene silencing in trans via H3K27 methylation and H3K4 demethylation. Recent investigation has found that HOTAIR can impose a further level of post-transcriptional regulation by acting as a competing endogenous RNA [60]. MiR-331-3p can directly bind HOTAIR, and HOTAIR functions as the endogenous decoy to disrupt the repression of HER2 by miR-331-3p. The increased HER2 expression may account for the more aggressive properties and poor survival associated with GC [60]. A case-control study in a northern Chinese population indicated that the T allele of rs12826786 might increase gastric cardia adenocarcinoma (GCA) risk, and this SNP has a genotype-specific influence on HOTAIR expression [88].

The oncogenic IncRNA ANRIL has also been found to be overexpressed in GC tissues, and higher ANRIL expression has been significantly correlated with an aggressive TNM stage and larger tumor size and has served as an independent predictor of poor overall survival. ANRIL exerts critical roles in cell proliferation both in vitro and in vivo [89]. The transcriptional activator E2F1 can induce ANRIL expression and promote ANRIL-mediated rapid cell growth. ANRIL can recruit and bind to the PRC2 complex to epigenetically repress miR-99a/miR-449a in trans, which activates the miR-99a/miR-449a target (the mTOR and CDK6/E2F1 pathway), thereby partially accounting for aberrant cell proliferation. Moreover, the formation of a positive feedback loop between ANRIL and E2F1 enables cell proliferation to be maintained continuously [89].

Overexpression of the imprinted IncRNA H19 gene is associated with GC development and poor prognosis. MiR-675 derived from H19 is a pivotal mediator in H19-induced GC development by silencing certain tumor suppressors. RUNX1 is a direct target of miR-675, and H19/miR-675-mediated RUNX1 depletion triggers cell proliferation and inhibits apoptosis in GC [75]. Independent of the miR-675 product, H19 has been found to interact with the tumor suppressor p53, abolishing p53 activity and thereby suppressing the expression of p53 targets, such as Bax, and leading to GC cell proliferation [90]. In addition, a coexpression network has revealed that ISM1 is the binding protein of H19, and its expression positively correlates with that of H19. CALN1 has been identified as another target of H19-derived miR-675. By directly upregulating ISM1 and indirectly repressing CALN1 expression by miR-675, H19 can promote cell proliferation, anti-apoptosis, migration, invasion and metastasis in GC. GC patients in the high H19 expression subgroup had shorter survival times [91]. H19 expression can be activated by the oncogene c-Myc in GC [90].

Multidrug resistance (MDR) is responsible for chemotherapy failure during GC treatment. The
**Long non-coding RNAs in gastric cancer**

IncRNA MRUL (MDR-related and upregulated IncRNA) has a significant effect on MDR. MRUL is located 400 kb downstream of ABCB1, which is an ATP-dependent efflux pump that eliminates toxic intracellular metabolic products. MRUL has been found to be significantly upregulated in adriamycin- or vincristine-resistant SGC7901 cells [92]. MRUL could positively impact ABCB1 expression in an orientation-and position-independent manner. MRUL depletion in multidrug-resistant cells enhanced the accumulation of adriamycin or vincristine, decreased adriamycin release, and promoted apoptosis in MDR GC cell lines. The high levels of MRUL in GC tissues were negatively correlated with the growth inhibition rates of GC specimens treated with chemotherapy drugs in vitro and predicted a poor prognosis [92].

The IncRNA GHET1 (gastric carcinoma high expressed transcript-1) is enhanced in GC tissues and is correlated with larger tumor size, increased tumor invasion and a poor survival rate. Gain-of-function and loss-of-function analyses have revealed that GHET1 can drive gastric carcinoma cell proliferation. A mechanistic investigation demonstrated that GHET1 physically interacted with the IGF2BP1 protein and facilitated the association of c-Myc mRNA with IGF2BP1, thus increasing the mRNA stability and expression of the oncogene c-Myc to promote proliferation [93].

The IncRNA CCAT1 has a higher expression in GC tissues than in normal counterparts, and ectopic expression of CCAT1 promotes cell proliferation and migration [94]. Computational screen and chromatin immunoprecipitation (ChIP) assays have confirmed that c-Myc binds directly to the E-box element within the promoter region of CCAT1, thereby increasing transcriptional activity of the CCAT1 promoter and CCAT1 levels. CCAT1 and c-Myc expression showed a strong correlation in GC [94].

MALAT1 was highly expressed in GC cell lines and induced SF2/ASF proteins to localize to the nucleolus. SF2/ASF is an important member of the serine/arginine-rich protein (SR) family and is involved in alternative splicing. MALAT1 depletion can arrest the cell cycle in the G0/G1 phase, inhibit proliferation, and impair the nuclear distribution and expression of SF2/ASF [95]. SF2/ASF-silencing can repress the cell cycle and proliferation, but SF2/ASF overexpression fails to rescue the effects induced by MALAT1 knockdown. These findings indicate that MALAT1 may modulate GC cell proliferation partly by regulating SF2/ASF expression and distribution [95].

The IncRNA HULC (highly upregulated in liver cancer) was originally identified in hepatocellular carcinoma (HCC) and acts as an oncogene to promote HCC development and progression [96]. Recently, HULC was shown to be significantly elevated in GC tissues and positively associated with lymph node metastasis, distant metastasis and advanced TNM stages. HULC overexpression enhances proliferation, invasion, and EMT and suppresses apoptosis in GC cells. Additionally, ectopic HULC expression could induce autophagy, whereas HULC silencing or treatment with an autophagic inhibitor increased apoptosis in SGC7901 GC cells. These results suggest that HULC might function as a molecular switch between autophagy and apoptosis during GC pathogenesis [97].

**The tumor suppressor roles of IncRNAs in GC**

The IncRNA FENDRR (FOXF1 adjacent non-coding developmental regulatory RNA) is an essential regulator of heart and body development in the mouse [98]. Similar to HOTAIR, FENDRR can bind to PRC2 to control chromatin structure and gene activity [99]. In GC tissues and cell lines, FENDRR expression has been shown to be lower, which is associated with tumor invasion, advanced tumor stage, lymphatic metastasis, and poor prognosis. Histone deacetylation is involved in the downregulation of FENDRR in GC cells [100]. The IncRNA FENDRR has no obvious effect on cell proliferation but plays important roles in cell invasion and migration. Forced FENDRR overexpression is able to effectively reduce the number of metastatic nodules in xenograft mouse models. Further investigation revealed that the reduction of fibronectin1 and MMP2/MMP9 expression accounted for the FENDRR-induced inhibition of GC cell metastasis [100].

The IncRNA GAS5 (growth arrest-specific transcript 5) has been identified as a tumor-suppressor IncRNA in bladder, pancreatic and breast cancer, as GAS5 is downregulated and exhibits the capacity for growth arrest in these cancers [101-103]. Sun et al. found that GAS5 exhibited decreased expression in GC tissues,
and the reduced expression of GAS5 was significantly associated with larger tumor size, advanced pathological stage, and poorer survival and overall survival. Increasing GAS5 expression was able to inhibit GC cell proliferation and induce apoptosis partly via the regulation of E2F1, P21, and cleaved caspase3 expression, although the mechanism is unclear \[104\]. Our ongoing investigations also demonstrated that GAS5 is downregulated in GC specimens, and GAS5 depletion triggers GC cell cycle arrest at the G1 phase. The results from the RNA-pull down and RNA immunoprecipitation (RIP) assays revealed that GAS5 interacted with the transcriptional activator YBX1 and regulated YBX1 protein abundance without affecting its mRNA level. The depletion of GAS5 and subsequent reduction in YBX1 protein decreased the expression of P21 transactivated by YBX1, eventually abolishing P21-induced cell cycle arrest in GC cells. Our findings suggest that the IncRNA GAS5/YBX1/p21 pathway plays critical roles in GC cell proliferation (Liu YC and Zhao J. et al. accepted).

The imprinted gene MEG3 (maternally expressed gene3) acts as a tumor suppressor IncRNA. MEG3 levels are markedly decreased in GC tissues and are correlated with an advanced TNM stage, increased invasion depth, larger tumor size and poor prognosis. DNA methylation is involved in MEG3 expression \[105\]. MEG3 overexpression can inhibit cell proliferation, drive apoptosis, and increase p53 expression in GC cells \[105\].

The IncRNA LincBM742401 is markedly downregulated in GC according to RNA-seq and public microarray data analyses. Reduced levels of BM742401 have been closely associated with poor survival in GC patients \[106\]. Forced expression of BM742401 in GC cells can block cell metastasis and decrease extracellular MMP9 secretion. Moreover, mice injected with BM742401-overexpressing cells had a significant reduction in tumor size and fewer lung metastatic foci \[106\].

The IncRNA ncRuPAR can increase protease-activated receptor-1 (PAR-1) levels during embryonic growth \[107\]. A recent study of GC tissues demonstrated that ncRuPAR was significantly downregulated, and the level of ncRuPAR was positively correlated with the tumor invasion depth, lymph node metastasis, distant metastasis, tumor size, and TNM stage. Moreover, decreased ncRuPAR expression was inversely associated with the mRNA level and immunohistochemical signal intensity of PAR-1 and VEGF in GC tissues \[107\], but the underlying mechanisms remain to be elucidated.

There is accumulating evidence that IncRNAs participate in GC development. A small fraction of IncRNAs have well-characterized mechanisms (Figure 1), but many dysregulated IncRNAs are not well understood and are correlated with clinical significance in GC. The GC-related IncRNAs have been summarized in Table 2.

The diagnosis and prognosis potential of IncRNAs in GCs

Early detection and diagnosis and effective prognostic indicators are necessary for improving the survival of GC patients. However, conventional tumor markers, such as CEA and CA
## Long non-coding RNAs in gastric cancer

### Table 2. The potential clinical applications of dysregulated lncRNAs in GC

<table>
<thead>
<tr>
<th>Oncogenic lncRNAs</th>
<th>Sample type and size</th>
<th>Methods</th>
<th>Clinical significance</th>
<th>Potential Application</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>H19↑</td>
<td>74 paired GC tissues</td>
<td>qRT-PCR</td>
<td>poor prognosis</td>
<td>diagnostic, prognostic biomarker</td>
<td>[91]</td>
</tr>
<tr>
<td></td>
<td>80 paired GC tissues</td>
<td>-</td>
<td>advanced TNM, poor prognosis</td>
<td>therapeutic target</td>
<td>[117]</td>
</tr>
<tr>
<td></td>
<td>24 paired GC tissues</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>[75]</td>
</tr>
<tr>
<td></td>
<td>22 paired GC tissues</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>[90]</td>
</tr>
<tr>
<td></td>
<td>43 pre-operative, 20 post-operative plasma, and 33 healthy plasma</td>
<td>Plasma H19 levels higher in GC patients and reduced in post-operative sample</td>
<td>-</td>
<td>-</td>
<td>[108]</td>
</tr>
<tr>
<td></td>
<td>77 paired GC tissues</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>[91]</td>
</tr>
<tr>
<td>HOTAIR↑</td>
<td>78 paired GC tissues</td>
<td>qRT-PCR</td>
<td>tumor size, advanced TNM, metastasis, and shorter survival</td>
<td>diagnostic, prognostic biomarker</td>
<td>[60]</td>
</tr>
<tr>
<td></td>
<td>68 paired GC tissues</td>
<td>-</td>
<td>venous invasion, lymph node metastases and lower survival</td>
<td>therapeutic target</td>
<td>[87]</td>
</tr>
<tr>
<td></td>
<td>31 paired GC tissues</td>
<td>-</td>
<td>TNM stage and lymph node metastasis</td>
<td>-</td>
<td>[115]</td>
</tr>
<tr>
<td></td>
<td>50 paired GC tissues</td>
<td>-</td>
<td>Lymphovascular invasion and lymph node metastasis</td>
<td>-</td>
<td>[86]</td>
</tr>
<tr>
<td></td>
<td>150 paired GC tissues</td>
<td>-</td>
<td>as independent prognostic and risk factor for peritoneal dissemination</td>
<td>-</td>
<td>[85]</td>
</tr>
<tr>
<td>GAPLINC↑</td>
<td>90 GC tissues</td>
<td>ISH</td>
<td>poor survival</td>
<td>diagnosis biomarker</td>
<td>[83]</td>
</tr>
<tr>
<td>MRUL↑</td>
<td>40 paired GC tissues, and SGC7901 cell line resistant VCR or ADR</td>
<td>qRT-PCR</td>
<td>in vitro growth inhibition rates of GC specimens treated with chemotherapy drugs and a poor prognosis</td>
<td>chemotherapy prediction and prognosis biomarker</td>
<td>[92]</td>
</tr>
<tr>
<td>ANRIL↑</td>
<td>120 paired GC tissues</td>
<td>qRT-PCR</td>
<td>TNM stage, tumor size and poor survival</td>
<td>diagnostic, prognostic biomarker and therapeutic target</td>
<td>[89]</td>
</tr>
<tr>
<td>GHET↑</td>
<td>42 paired GC tissues</td>
<td>qRT-PCR</td>
<td>tumor size, tumor invasion and poor survival</td>
<td>diagnostic, prognostic biomarker and therapeutic target</td>
<td>[93]</td>
</tr>
<tr>
<td>CCAT↑</td>
<td>20 paired GC tissues</td>
<td>qRT-PCR</td>
<td>-</td>
<td>therapeutic target</td>
<td>[94]</td>
</tr>
<tr>
<td>HULC↑</td>
<td>58 paired GC tissues</td>
<td>qRT-PCR</td>
<td>lymph node metastasis, distant metastasis and advanced TNM stages</td>
<td>diagnosis biomarker and therapeutic target</td>
<td>[97]</td>
</tr>
<tr>
<td>MALAT1↑</td>
<td>150 paired GC tissues</td>
<td>qRT-PCR</td>
<td>correlated with peritoneal metastasis in GC patients</td>
<td>diagnosis and prognosis biomarker, and therapeutic target</td>
<td>[85]</td>
</tr>
<tr>
<td>Linc00152↑</td>
<td>79 GC plasma, and paired pre- and post-operative plasma</td>
<td>qRT-PCR</td>
<td>Linc00152 levels higher in GC patients plasmas</td>
<td>diagnosis as blood-based biomarker</td>
<td>[110]</td>
</tr>
<tr>
<td></td>
<td>71 paired GC tissues, 17 gastric juice of GC and 16 normal mucosa or minimal gastritis</td>
<td>-</td>
<td>correlated with invasion, diagnostic potential</td>
<td>-</td>
<td>[109]</td>
</tr>
<tr>
<td>FENDR↓</td>
<td>158 paired GC tissues</td>
<td>qRT-PCR</td>
<td>tumor invasion, TNM, lymphatic metastasis and poor prognosis</td>
<td>diagnosis biomarker</td>
<td>[121]</td>
</tr>
<tr>
<td>GASS↓</td>
<td>89 paired GC tissues</td>
<td>qRT-PCR</td>
<td>tumor size, advanced TNM and prognosis</td>
<td>diagnostic, prognostic biomarker</td>
<td>[104]</td>
</tr>
<tr>
<td>ncRuPAR↓</td>
<td>55 paired GC tissues</td>
<td>qRT-PCR</td>
<td>tumor invasion depth, lymph node metastasis, distant metastasis, tumor size, and TNM stage</td>
<td>diagnostic, prognostic biomarker</td>
<td>[107]</td>
</tr>
<tr>
<td>uc001lsz↓</td>
<td>138 paired GC tissues</td>
<td>qRT-PCR</td>
<td>TNM stage</td>
<td>early diagnosis marker</td>
<td>[20]</td>
</tr>
<tr>
<td>MEG3↓</td>
<td>72 paired GC tissues</td>
<td>qRT-PCR</td>
<td>TNM stages, depth of Invasion, tumor size and poor prognosis</td>
<td>prognosis biomarker</td>
<td>[105]</td>
</tr>
<tr>
<td>Tumor suppressor lncRNAs</td>
<td>Sample type and size</td>
<td>Methods</td>
<td>Clinical significance</td>
<td>Potential Application</td>
<td>Ref</td>
</tr>
<tr>
<td>FENDR↓</td>
<td>158 paired GC tissues</td>
<td>qRT-PCR</td>
<td>tumor invasion, TNM, lymphatic metastasis and poor prognosis</td>
<td>diagnosis biomarker</td>
<td>[100]</td>
</tr>
<tr>
<td>GASS↓</td>
<td>89 paired GC tissues</td>
<td>qRT-PCR</td>
<td>tumor size, advanced TNM and prognosis</td>
<td>diagnosis biomarker and therapeutic target</td>
<td>[104]</td>
</tr>
<tr>
<td>ncRuPAR↓</td>
<td>55 paired GC tissues</td>
<td>qRT-PCR</td>
<td>tumor invasion depth, lymph node metastasis, distant metastasis, tumor size, and TNM stage</td>
<td>diagnosis biomarker</td>
<td>[107]</td>
</tr>
<tr>
<td>uc001lsz↓</td>
<td>77 paired GC tissues</td>
<td>qRT-PCR</td>
<td>TNM stage</td>
<td>early diagnosis marker</td>
<td>[20]</td>
</tr>
<tr>
<td>MEG3↓</td>
<td>72 paired GC tissues</td>
<td>qRT-PCR</td>
<td>TNM stages, depth of Invasion, tumor size and poor prognosis</td>
<td>diagnosis biomarker</td>
<td>[105]</td>
</tr>
</tbody>
</table>
## Long non-coding RNAs in gastric cancer

<table>
<thead>
<tr>
<th>LncRNAs SNP</th>
<th>Sample type and size</th>
<th>Methods</th>
<th>Clinical significance</th>
<th>Potential Application</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>HOTAIR</td>
<td>515 GCA and 654 healthy control blood</td>
<td>PCR-RFLP; qRT-PCR</td>
<td>T allele of rs12826786 was risk for GCA and associated with advanced TNM</td>
<td>Screening and prognosis biomarker</td>
<td>[88]</td>
</tr>
<tr>
<td>CAS08</td>
<td>940 GC tissues</td>
<td>SNaP shot</td>
<td>GG genotype of rs10505477 survived for a longer time</td>
<td>Prognostic biomarker</td>
<td>[113]</td>
</tr>
</tbody>
</table>

†: upregulated; ↓: downregulated; GC: gastric cancer; GCA: gastric cardia adenocarcinoma; PCR-RFLP: polymerase chain reaction-restriction fragment length polymorphism; qRT-PCR: quantitative real-time reverse transcription PCR; ISH: in situ hybridization.

<table>
<thead>
<tr>
<th>BM742401</th>
<th>113 paired GC tissues</th>
<th>qRT-PCR</th>
<th>poor survival</th>
<th>Prognostic biomarker and therapeutic target</th>
<th>[106]</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC096655.1-002 (GACAT1)</td>
<td>78 paired GC tissues</td>
<td>qRT-PCR</td>
<td>lymph node metastasis, distant metastasis, TNM, and differentiation</td>
<td>Diagnosis biomarker</td>
<td>[124]</td>
</tr>
<tr>
<td>AA174084</td>
<td>134 paired GC tissues, 127 gastric mucosal tissues, 335 plasma, and 130 gastric juice</td>
<td>qRT-PCR</td>
<td>Borrmann type, perineural invasion and lymphatic metastasis, dropped in post-operation</td>
<td>Early diagnosis biomarker</td>
<td>[111]</td>
</tr>
<tr>
<td>FER1L4</td>
<td>61 paired GC tissues, 80 healthy control plasma and 83 paired pre- and postoperative plasma</td>
<td>qRT-PCR</td>
<td>tumor size, histologic grade, general classification, depth of invasion, lymphatic metastasis, distant metastasis, TNM stage, vessel or nerve invasion</td>
<td>Prognostic biomarker</td>
<td>[125]</td>
</tr>
<tr>
<td>HMlincRNA717 (GACAT2)</td>
<td>107 paired GC tissues, 37 healthy gastric mucosa, 34 gastritis mucosa, and 28 gastric precancerous lesions</td>
<td>qRT-PCR</td>
<td>cancer distal metastasis, venous invasion and nervous invasion</td>
<td>Early diagnosis biomarker</td>
<td>[126]</td>
</tr>
<tr>
<td>AC138128.1</td>
<td>94 paired GC tissues</td>
<td>qRT-PCR</td>
<td>-</td>
<td>Diagnosis biomarker</td>
<td>[127]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Sample type and size</th>
<th>Methods</th>
<th>Clinical significance</th>
<th>Potential Application</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>515 GCA and 654 healthy control blood</td>
<td>PCR-RFLP; qRT-PCR</td>
<td>T allele of rs12826786 was risk for GCA and associated with advanced TNM</td>
<td>Screening and prognosis biomarker</td>
<td>[88]</td>
</tr>
<tr>
<td>940 GC tissues</td>
<td>SNaP shot</td>
<td>GG genotype of rs10505477 survived for a longer time</td>
<td>Prognostic biomarker</td>
<td>[113]</td>
</tr>
</tbody>
</table>
Long non-coding RNAs in gastric cancer

19-9, have limited sensitivity and specificity in routine screening for GC. Therefore, novel biomarkers are urgently needed for GC. As IncRNAs play versatile roles in the regulation of gene expression via transcription, post-transcription and especially epigenetic modulations, aberrant IncRNA expression may therefore occur during carcinogenesis and disease progression, and dysregulated IncRNA levels may be better indicators of the intrinsic properties of cancer. Furthermore, IncRNAs have tissue-specific expression patterns and can be detected in body fluids. These advantages make IncRNAs promising biomarkers for diagnosis, prognosis and therapy in various cancers, including GC.

The diagnostic potential of IncRNAs in GC

The ideal biomarker should be easily and non-invasively accessible. Test samples from body fluids, such as blood and gastric juice, are good choices. A previous study evaluated levels of the IncRNAs H19, HOTAIR and MALAT1 in the plasma of GC patients, and H19 was markedly stable and could be successfully amplified with qRT-PCR, suggesting the potential of evaluating circulating IncRNAs as biomarkers. The authors found that the H19 plasma levels were significantly higher in 43 GC patients than in 34 healthy controls [108]. Moreover, H19 levels were markedly lower in 16 one-month post-operative patients than in the pre-operative plasma from those same 16 patients, indicating that the circulating IncRNAs might be released from primary GCs [108]. These findings suggest that circulating IncRNAs, such as H19, could become new complementary tumor biomarkers for GC.

The IncRNA linc00152 is another GC-related IncRNA that is significantly upregulated in GC tissues compared with adjacent normal tissues [109]. The linc00152 plasma levels in 79 GC patients were significantly higher than those in 81 healthy controls. The receiver operating characteristic (ROC) curve assessing the diagnostic value of plasma linc00152 showed that the area under the ROC curve reached 0.657 (p < 0.001), with a sensitivity of 48.1% and a specificity of 85.2%. It is interesting to note that the sensitivity of linc00152 was better than that of the traditional markers CEA and CA19-9 [110]. This study also indicated that the stable existence of IncRNAs in the blood might result from exosome protection, as there were also no differences in the linc00152 levels between the plasma and exosomes isolated from the same plasma [110]. These results suggest that plasma linc00152 has great potential for GC diagnosis. Recently, levels of the IncRNA AA174084 were assessed in 134 paired GC tissues, 127 gastric mucosal tissues, 335 plasma samples, and 130 gastric juice samples at various stages of GC [111]. It was found that AA174084 was significantly reduced in 95 of 134 GC tissues (71%) compared with paired normal samples, and the reduced expression of AA174084 was negatively associated with Bormann type GC and perineural invasion [111]. The plasma AA174084 levels decreased significantly on day 15 after surgery compared with preoperative samples and were correlated with invasion and lymphatic metastasis [111]. Furthermore, AA174084 could be detected in gastric juice by qRT-PCR, and gastric juice from GC patients had significantly higher AA174084 levels than those from the patients with normal mucosa or with minimal gastritis, gastric ulcers, and atrophic gastritis. The area under the ROC curve was up to 0.848 (p < .001), with a sensitivity of 46% and a specificity of 93%, which were higher than those obtained using the tissue level of AA174084 as the biomarker. These data demonstrated that the AA174084 gastric juice level may be a potential screening biomarker for the early diagnosis of GC [111].

Recently, Guo et al. performed a case-control study in northern China to evaluate the association between haplotype-tagging SNPs (htSNPs) of HOTAIR and the susceptibility to gastric cardia adenocarcinoma (GCA). PCR-RFLP was used to detect the htSNP genotype for HOTAIR (rs12826786 C > T, rs4759314 A > G, and rs10783618 C > T) in 515 GCA patients (blood from before surgery) and 654 healthy controls. They found that the T allele of rs12826786 increased the susceptibility to GCA and was associated with smoking and the TNM stage. qRT-PCR results revealed that the rs12826786 SNP had a genotype-specific effect on HOTAIR expression. This study indicated that the rs12826786 SNP of HOTAIR could be a useful candidate biomarker for high-risk GCA populations [88].

There are other dysregulated IncRNAs in GC tissues that have the potential to be candidate diagnostic biomarkers, including GACAT1, ncrRuPAR, HULC, and FER1L4, among others.
Long non-coding RNAs in gastric cancer

(see Table 2). Whether the altered expression of these IncRNAs can be validated in body fluids remains to be elucidated.

The prognostic potential of IncRNAs in GC

Great advances have been made in IncRNA-based prognosis biomarker research. For instance, the prostate-specific IncRNA PCA3 has become the first FDA-approved IncRNA-based cancer biomarker to predict the prognosis of prostate cancer [112]. In GC, there are many dysregulated IncRNAs that are closely associated with a poor prognosis, and these IncRNAs could become desirable candidates for monitoring high-risk populations and predicting GC outcomes. An effective prognosis prediction may help GC patients gain access to more appropriate treatments.

The high cancer risk gene desert region 8q24 has a genetic variant SNP rs10505477 located within the intron of the IncRNA CASC8 that may affect the folding structures of CASC8. Ma et al. found that GC patients with rs10505477 GG survived for a longer time compared with those carrying the GA and AA genotypes in 940 surgical GC specimens [113]. This risk effect was more significant among patients with a tumor size ≤ 5 cm, diffuse-type GC, lymph node metastasis, no distant metastasis and TNM stage III and IV. These findings suggest that SNP rs10505477 may be a potential prognostic biomarker for GC [113].

MALAT1 and HOTAIR were expressed in 150 GC patients at higher levels than in corresponding adjacent normal mucosa and were markedly correlated with peritoneal metastasis in GC patients [85]. Additionally, HOTAIR overexpression was not only an independent prognostic indicator but also a risk factor for peritoneal dissemination [85]. Lee et al. reported that higher levels of HOTAIR were associated with lymphovascular invasion, lymph node metastasis, advanced TNM stage, and inferior disease-free survival in 50 GC samples [86]. Increased HOTAIR expression has been further confirmed to be positively correlated with an aggressive clinical significance in GC patients in many independent studies [60, 87, 114-116]. H19 is another extensively studied IncRNA with the potential to be a prognostic biomarker. High levels of MRUL in GC tissues were negatively associated with rates of growth inhibition in GC specimens treated with the chemotherapeutic drugs adriamycin or vincristine in vitro [92]. Furthermore, upregulated MRUL indicated a poor prognosis for GC patients. Thus, MRUL is a potential target for reversing multidrug-resistance in GC [92].

Variations in the IncRNAs GAPLINC, ANRIL, GHET1, HULC GAS5, MEG3, BM742401, and ncRuPAR have also been found to be significantly correlated with an aggressive TNM stage, increased invasion depth, lymph node metastasis, distant metastasis, and poorer survival rate (Table 2).

In addition, certain IncRNAs that are specifically overexpressed in cancer tissues, such as H19, are promising therapeutic targets for decreasing the “off-target” effects of gene therapy for cancers. Plasmids containing diphtheria toxin subunit A under the H19 promoter have been studied for curing colon, pancreatic, bladder and ovarian cancers via intratumoral injection [118], as the high H19 levels in cancer cells could induce subunit A expression to kill cancer cells while simultaneously protecting normal cells from this destruction. H19 is upregulated in GC, and the effect of the H19 vector on GC therapy should be explored.

Despite some progress, the clinical translation of IncRNAs in GC remains in its infancy. There are still many puzzles that must be solved via elaborate investigations. For example, cancer stem cells played important roles in cancer progression in our previous studies [119, 120], and it remains to be determined whether IncRNAs are involved in the maintenance of cancer stem cells in GC. In addition, it is unclear whether the circulating IncRNA alterations in the body fluids of GC patients account for GC development and if they are secreted from the GC cells or from other cell types. The stability of IncRNAs as biomarkers remains largely unknown but should be addressed to determine the general utility of IncRNAs in clinical practice.

Conclusions

IncRNAs can regulate gene expression at the transcriptional, post-transcriptional, and epigenetic architectures and play important roles in carcinogenesis and aggressive progression.
Many dysregulated lncRNAs in GC have been significantly associated with increased tumor size, invasion and metastasis, and poor survival outcomes. Certain dysregulated lncRNAs are promising candidate molecular biomarkers in GC. Future studies are needed to identify additional cancer-specific lncRNAs, validate the utility of lncRNA-based biomarkers for diagnosis and prognosis, and address the roles of lncRNAs in cancer biology, as well determine whether lncRNAs have the potential to serve as improved therapeutic targets.

Acknowledgements

This study was supported by the Natural Science Foundation of China (81201521), the China Postdoctoral Science Foundation Funded Project (2013M541467), and the Foundation from Science and Technology Commission of Shanghai (13XD1401200).

Disclosure of conflict of interest

None.

Address correspondence to: Wen-Hong Zhang or Ying Zhang, Department of Infectious Diseases, Huashan Hospital, Shanghai Medical College, Fudan University, Shanghai 200040, China. E-mail: zhan-gwenhong@fudan.edu.cn (WHZ); yzhang@jhsp.edu (YZ)

References

cceptibility identified using a genome-wide as-
[7] Macdonald JS, Smallery SR, Benedetti J, Hund-
dahl SA, Estes NC, Steermann GN, Haller DG, Ajani JA, Gunderson LL, Jessup JM and Martenson JA. Chemoradiotherapy after sur-
gery compared with surgery alone for adeno-
carcinoma of the stomach or gastroesophage-
[8] De Angelis R, Sant M, Coleman MP, Francis-
sci S, Baili P, Pierannunzio D, Trama A, Visser O, Brenner H, Ardazan E, Bielska-Lasota M, Eng-
holm G, Nennecke A, Siesling S, Berrino F and Capocaccia R. Cancer survival in Europe 1999-
2007 by country and age: results of EU-
ROCARE–5-a population-based study. Lancet Onc-
genome reconstruction and mutual signa-
zweig ER, Karnani N, Lee K, Lefebvre GC, Na-
ford GE, Sunyaev S, Noble WS, Dunham I, De-
Long non-coding RNAs in gastric cancer

Long non-coding RNAs in gastric cancer


Long non-coding RNAs in gastric cancer


[68] Faghihi MA, Morarresi F, Khalil AM, Wood DE, Sahagan BG, Morgan TE, Finch CE, St Laurent G 3rd, Kenny PJ and Wahlestedt C. Expression of a noncoding RNA is elevated in Alzheimer’s disease and drives rapid feed-forward regula-
Long non-coding RNAs in gastric cancer

mRNA is a primary microRNA precursor. RNA 2007; 13: 313-316.


Wang Y, Zhang D, Wu K, Zhao Q, Nie Y and Fan D. Long non-coding RNA MRUL promotes ABCB1 expression in multidrug-resistant gas-
Long non-coding RNAs in gastric cancer


